Oszillationstheorem für matrixwertige Sturm-Liouville-Probleme

Diplomarbeit
von
Martina Comes

Institut für Mathematik
der Universität Erlangen-Nürnberg

Lehrstuhl für Mathematische Physik

Erlangen 2007

Inhaltsverzeichnis

Kapitel 1

Einleitung

Unter allen Disziplinen der Mathematik ist die *Theorie der Differentialgleichungen* die wichtigste. Alle Zweige der Physik stellen uns Probleme, die auf die Integration von Differentialgleichungen hinauskommen. Es gibt die Theorie der Differentialgleichungen den Weg zur Erklärung aller elementarer Naturphänomene, die Zeit brauchen.

Hat somit die Theorie der Differentialgleichungen eine unendliche *praktische* Bedeutung, so hat sie auf der anderen Seite dadurch eine entsprechende *theoretische* Wichtigkeit, dass sie in rationeller Weise zum Studium neuer wichtiger Funktionen beziehungsweise Funktionenklassen leitet.

Sophus Lie (1894), zitiert nach [Am95, S. V.].

Spricht Sophus Lie von der *„unendliche(n) praktische(n) Bedeutung“* von Differentialgleichungen, so scheint das zunächst maßlos übertrieben. Doch man bedenke: Ob die schwingende Saite oder ein Pendel, ob die Bewegung von Himmelskörpern oder das Wachstum von Muskeln, ob die Populationsdynamik oder die Black-Scholes-Formel zur Bewertung von Finanzoptionen – in all diesen Fällen werden Differentialgleichungen genutzt, um Probleme zu beschreiben und zu lösen. Insofern ist der erste Teil des obigen Zitats gerechtfertigt.

In dieser Arbeit steht das *Sturm-Liouville-Problem* im Fokus, ein Differential-Operator zweiten Grades, der in dieser Arbeit mit $\mathcal{L}$ bezeichnet wird. Das Ziel der Betrachtungen ist es, Lösungen der Gleichung $\mathcal{L}y = Ey$ zu finden. Die Sturm-Liouville-Gleichungen gehören zu den bedeutendsten Differentialgleichungen, die es zu lösen gilt. Die Sturm-Liouville-Theorie wird eingesetzt, um Probleme wie Wärmeverteilung und Schwingungen zu modellieren, und sie spielt eine wichtige Rolle in der Quantenmechanik, wo die auftretenden Probleme allerdings meist singulär sind.

1.1 Historischer Überblick

Die Suche nach den Lösungen des Sturm-Liouville-Problems beschäftigt die Wissenschaft schon sehr lange. Erstmals aufgeworfen wurde das Problem in einer Reihe von Veröffentlichungen, die in den 1830er Jahren erschien, und in der der schweizer Mathematiker *Charles Sturm* (1803–1855) und der französische Mathematiker *Joseph Liouville* (1809–1882) die qualitativen Eigenschaften der folgenden Differentialgleichung studierten:

$$\frac{\mathrm{d}}{\mathrm{dx}}\left(K\,\frac{\mathrm{d}V}{\mathrm{dx}}\right)+GV=0 \qquad\qquad \text{für } x\geq\alpha.$$

Dabei waren K, G und V reelle Funktionen einer Variablen x. Das Ziel ihrer Untersuchungen war, aus der Form der Gleichungen sofort Eigenschaften der Lösungen herzuleiten – ohne die Lösungen zu kennen (vgl. [Hin05]). Damit gaben die beiden den Anstoß zur Entwicklung der Sturm-Liouville-Theorie, die sich mit der Existenz und dem asymptotischen Verhalten der entsprechenden Eigenwerte befasst. Dies bedeutete einen ersten Schritt in Richtung einer Spektraltheorie für Differentialoperatoren zweiten Grades. In moderner Notation lautet obige Gleichung:

$$-(p_x y_x')'+v_x y_x=0, \qquad\qquad x\in[\alpha,\,+\infty).$$

Dabei sind p_x und v_x reelle Funktionen, deren Differenzier- und Integrierbarkeit je nach Problemstellung festgelegt wird.

Im 20. Jahrhundert kam es zu großen Fortschritten bei der Betrachtung des singulären Problems (speziell für das Intervall $I=[0,\,\infty)$ bzw. den Hilbertraum $L^2[0,\,\infty)$).[1] 1910 veröffentlichte Hermann Weyl einen Aufsatz, der die Spektraltheorie für Sturm-Liouville-Probleme geradezu revolutionierte ([Wey10]). Insbesondere zeigte er, dass man singuläre Sturm-Liouville-Differentialoperatoren in zwei Gattungen aufteilen kann: In einen *limit-point-* und den *limit-circle-Fall* (vgl. beispielsweise [Sa99, Kap. 9]). Auch von Edward Charles Titchmarsh und Kunihiko Kodaira wurde die Theorie später ausgebaut. Eine ausführliche Darstellung der historischen Entwicklung findet man in [Ev05a, Hin05]. Insofern führte die Beschäftigung mit der von Sturm und Liouville formulierten Differentialgleichung in der Tat zu theoretischen Neuerungen.

[1]Details zu den Entwicklungen von Weyl und seinen Nachfolgern findet man in [Ev05b].

1.2 Hauptresultate und Aufbau dieser Diplomarbeit

In dieser Arbeit werden die Lösungen *regulärer* Sturm-Liouville-Probleme diskutiert. Es wird sich zeigen, dass auch die Beschäftigung mit dieser Problemgattung zu interessanten Resultaten führt.

Definition 1.1 *Der hier diskutierte Operator hat die Form:*

$$(\mathcal{L}y)_x \equiv -\big(p_x y'_x + q_x y_x\big)' + q^*_x y'_x + v_x y_x, \qquad x \in [0,\ 1]$$

und wird auf dem Bereich

$$\mathcal{D}(\mathcal{L}) = \left\{ y \in C^1\left(I, \mathbb{C}^L\right) :\ y' \in AC\left(I, \mathbb{C}^L\right),\ u_j \in \Psi_j \mathbb{C}^L,\ (j = 0,\ 1) \right\}$$

betrachtet. Dabei sind p_x, q_x und $v_x \in \mathrm{Mat}(L \times L,\ \mathbb{C})$, p_x und q_x absolut stetig sowie v_x stetig auf $[0,\ 1]$. $u_j \in \mathbb{C}^{2L}$ ist gegeben durch

$$u_j = \begin{pmatrix} y_j \\ p_j y'_j + q_j y_j \end{pmatrix}, \qquad (j = 0,\ 1).$$

Außerdem ist $\Psi_j \in \mathrm{Mat}(2L \times L,\ \mathbb{C})$ derart, dass stets die Bedingung

$$\Psi^*_j \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix} \Psi_j = 0 \qquad (j = 0,\ 1).$$

erfüllt ist. Dirichlet-Randbedingungen entsprechen:

$$u_0 \in \begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix} \mathbb{C}^L, \quad u_1 \in \begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix} \mathbb{C}^L. \tag{1.1}$$

Man untersucht nun stets das Sturm-Liouville-Problem

$$\mathcal{L}y = Ey. \tag{1.2}$$

Im Fall $L = 1$ sind folgende Resultate bekannt:

1. **Existenzsatz:** Es gibt zu dem Eigenwertproblem (1.2) unendlich viele reelle Eigenwerte, vgl. [Wa93, Kap. 27, S. 239].

2. **Oszillationstheorem:** Die Lösungen des Problems (1.2) oszillieren. Eine Lösung y oszilliert genau dann, wenn y abzählbar viele Nullstellen auf dem kompakten Intervall I hat, vgl. [Wa93, Kap. 27, S. 246]. Es gilt: die Eigenwerte λ_0, λ_1, ... bilden eine monoton wachsende Folge, wobei $\lambda_j \to \infty$ für $j \to \infty$. Darüber hinaus hat die Eigenfunktion zum Eigenwert λ_j genau j Nullstellen im Intervall $[0,\ 1]$, vgl. [CL55, Kap. 8, S. 212]. Das heißt, zu jedem festen $E \in \mathbb{R}$ gibt es genau so viele Nullstellen (d. h. Oszillationen) wie Eigenwerte $\leq E$ vorliegen.

Für matrixwertige Funktionen sind diese Resultate wesentlich schwieriger zu bestimmen als im eindimensionalen Fall. In den 1950er Jahren veröffentlichte von Raoul Bott die Arbeit [Bo56], in der er ein Oszillationstheorem für matrixwertige Sturm-Liouville-Probleme beweist. Eine ähnliche Methode wie Bott wendet auch diese Arbeit an. Es ist aber gelungen, seine Vorgehensweise wesentlich zu vereinfachen. Außerdem werden genauere Aussagen über das Verhalten für $E \to -\infty$ gemacht, als sie bei Bott zu finden sind.

Um die Hauptresultate dieser Arbeit zusammenzufassen benötigt man die folgende kurze Definition:

Definition 1.2 *Es sei* $\Phi_x^E = \begin{pmatrix} a_x^E \\ b_x^E \end{pmatrix} \in \mathrm{Mat}(2L \times L,\ \mathbb{C})$ *eine Lösung des Sturm-Liouville-Problems an einem Punkt* $x \in [0,\ 1]$ *zu einer vorgegebenen Anfangsbedingung. Dann definiert man:*

$$U_x^E = \left(a_x^E - ib_x^E\right)\left(a_x^E + ib_x^E\right)^{-1}.$$

Satz 1.1 (Hauptresultate)
Die folgenden Resultate gelten für den in Definition 1.1 vorgestellten Operator $\mathcal{L}$ *auf dem Definitionsbereich* $\mathcal{D}(\mathcal{L})$ *und unter Dirichlet-Randbedingungen* (1.1).

1. U_1^E *ist wohldefiniert, d. h. die Inverse* $\left(a_1^E + ib_1^E\right)^{-1}$ *existiert.*

2. *Es gilt:* U_1^E *ist unitär und analytisch in* E.

3. *Sind* $e^{i\theta_l^E}$ *die Eigenphasen von* U_1^E $(l = 1,\ \ldots,\ L)$, *so kann jede Funktion* θ_l^E *analytisch in* E *gewählt werden.*

4. *Es gilt:*

$$\begin{aligned} & \textit{Multiplizität von } E \textit{ als Eigenwert des Sturm-Liouville-Problems} \\ = \quad & \textit{Multiplizität von } e^{i\pi} \textit{ als Eigenwert von } U_1^E. \end{aligned}$$

5. *Die Matrix* $S_1^E := \frac{1}{i}\left(U_1^E\right)^* \partial_E U_1^E$ *ist positiv definit und alle* θ_l^E *sind monoton wachsend in* E.

6. *Falls* $\mathcal{L}$ *nach unten beschränkt und* $q_x \equiv 0$ *ist, können die Funktionen* θ_l^E *(*$l = 1, \ldots, L$*) derart gewählt werden, dass gilt:*

$$\theta_l^E \to -\pi \qquad \text{für } E \to -\infty.$$

Wie gelangt man zu diesen Resultaten?

Zunächst wird die Problemstellung näher umrissen, daher steht die Beschäftigung mit dem *regulären Sturm-Liouville-Operator* $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ und seinen Eigenschaften (beispielsweise Selbstadjungiertheit) im Vordergrund des zweiten Kapitels. Weiterhin wird das Problem umformuliert: Es wird gezeigt, dass das Problem $\mathcal{L}y = Ey$ äquivalent zu einem System erster Ordnung ist.

Das dritte Kapitel ist der *hermitesch symplektischen Geometrie* gewidmet. Zunächst wird eine Einführung in diese Geometrie und ihre Eigenschaften gegeben, und man definiert die hermitesch symplektische Gruppe, $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$, sowie die zugehörige hermitesch symplektische Lie-Algebra, $\mathfrak{sp}(2L, \mathbb{C})$. Danach wird gezeigt, dass die Lösungen des Sturm-Liouville-Problems Lagrange-Ebenen sind. Weiterhin kann man zeigen, dass man die Pfade der Lösungen mit Hilfe einer hermitesch symplektischen Matrix beschreiben kann, die auf die Lösungsebene am linken Rand angewendet wird.

Das vierte Kapitel beweist die Hauptresultate 1–3. Dazu werden die *stereographische Projektion* und die *Cayley-Transformation* $\mathcal{C}$ zunächst detailliert erläutert. Anschließend wird gezeigt, dass eine Verknüpfung dieser beiden Abbildungen, die mit Π bezeichnet wird, die Lagrange-Ebenen auf die unitären Matrizen abbildet. Damit ist ein Diffeomorphismus zwischen den Lösungen des Sturm-Liouville-Problems und den unitären Matrizen etabliert. Weiterhin definiert man die *Lorentz-Gruppe* als diejenige Gruppe, die durch die Cayley-Transformation der hermitesch symplektischen Gruppe entsteht: $\mathcal{U}(L, L, \mathbb{C}) = \mathcal{C}\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})\mathcal{C}^*$.

Im fünften Kapitel beschäftigen wir uns intensiv mit den unter Π entstehenden Bildern. Es wird gezeigt, dass man den maximalen Rand der Einheitskreise mit der unitären Gruppe identifizieren kann. Daher etabliert Π einen Diffeomorphismus zwischen den Lagrange-Ebenen und dem maximalen Rand des Einheitskreises. Um das Verhalten der Lösungen in der Zeit auf dem Raum der unitären Matrizen beschreiben zu können, ist ein weiterer Abschnitt des fünften Kapitels der *Möbius-Transformation* gewidmet. Damit ist man in der Lage, die Differentialgleichung mittels Π in eine Differentialgleichung für unitäre Matrizen umzuformulieren.

Das sechste Kapitel beschäftigt sich mit Strukturen, die es erlauben, einen Schnittindex zu definieren. Dazu definiert man singuläre Zykel. Mit Hilfe der *Wronski-Matrix* gelingt es, eine Aussage über die Dimension der Schnitte zu machen. Weiterhin wird der *Bott-Maslov-Index* beschrieben. Dieser Index erlaubt es, für geschlossene Pfade einen Schnittindex explizit zu berechnen.

Im siebten Kapitel werden die Resultate zusammengefügt, und man erhält die Aussagen 4 und 5.

Im achten Kapitel steht das Konvergenzverhalten von U_1^E für $E \to -\infty$ im Mittelpunkt. Mittels eines Homotopiearguments und der 5. Hauptaussage gelingt es, Hauptresultat 6 zu beweisen.

Darüber hinaus wird in Anhang B gezeigt, dass im Falle $\mathcal{L} = \overline{\mathcal{L}}$ (d. h. im Falle eines reellen Sturm-Liouville-Operators) gilt:

1. U_1^E ist symmetrisch.
2. S_1^E ist reell.

Ein analoges Ergebnis wird auch für den quaternionischen Fall hergeleitet.

Kapitel 2

Reguläre Sturm-Liouville-Probleme auf kompakten Intervallen

Der Beginn aller Wissenschaften ist das Erstaunen,
dass die Dinge sind, wie sie sind.

Aristoteles, Metaphysik.

2.1 Präliminarien

Sei $\mathbb{C}^L$ der Raum der komplexen L-Tupel mit Skalarprodukt, in Dirac-Notation: $\langle\alpha|\beta\rangle = \sum_{i=1}^{L} \overline{\alpha_i}\,\beta_i$, wobei $\overline{\alpha}$ das komplex Konjugierte einer komplexen Zahl α sei. Weiterhin bezeichne p^* die adjungierte Abbildung zu einer linearen Abbildung $p : \mathbb{C}^L \to \mathbb{C}^L$. Somit gilt $\langle p\,\alpha|\beta\rangle = \langle\alpha|p^*\beta\rangle$. In einem Koordinatensystem des $\mathbb{C}^L$ entspricht die Matrixdarstellung der Abbildung p^* der Matrix $\overline{p}^{\,T}$, wobei unter p^T stets die transponierte Matrix zu verstehen ist.

Um die in dieser Arbeit untersuchten Zusammenhänge zu beschreiben, erweist es sich als sinnvoll, $\mathbb{C}^{2L}$ als $\mathbb{C}^L \oplus \mathbb{C}^L$ aufzufassen. Das Skalarprodukt in $\mathbb{C}^{2L}$ wird folgendermaßen aus dem Skalarprodukt des $\mathbb{C}^L$ abgeleitet: Sei $u = \{\alpha_1,\ \beta_1\}$ und $v = \{\alpha_2,\ \beta_2\}$ (mit $\alpha_i,\ \beta_i \in \mathbb{C}^L$ $(i = 1,\ 2)$), dann ist $\langle u|v\rangle = \langle\alpha_1|\alpha_2\rangle + \langle\beta_1|\beta_2\rangle$.

Notationshinweise: Im gesamten Text bezeichne $y' = y'(x)$ die Ableitung von $y(x)$ nach x. Auf die Abhängigkeit von x wird – aus Gründen der Übersichtlichkeit – immer mit einer Indizierung hingewiesen, d. h. $y(x) =: y_x$. Außerdem stehe $\mathbb{1}_L$ für die $L \times L$ -Einheitsmatrix. Auf die Indizierung wird verzichtet, wenn die Dimension aus dem Zusammenhang klar hervorgeht.

2.2 Reguläre Sturm-Liouville-Operatoren auf kompakten Intervallen

In der Literatur findet man – abhängig vom Ziel der Untersuchungen – verschiedene Darstellungen und Formen der Sturm-Liouville-Operatoren. Die Gestalt des Operators, der in dieser Arbeit im Mittelpunkt steht, ist in Anlehnung an die Arbeit von Bott [Bo56] gewählt. Da diese Arbeit sich nur mit dem regulären Sturm-Liouville-Problem beschäftigt sei stets $I = [0,\, 1]$ und $x \in I$.[1]

Um den Sturm-Liouville-Operator $\mathcal{L}$ auf einem geeigneten Bereich $\mathcal{D}(\mathcal{L})$ zu definieren, müssen einige Definitionen vorausgeschickt werden. Zunächst benötigt man folgende Definition, mit deren Hilfe es möglich sein wird, die Randbedingungen von $\mathcal{L}$ zu beschreiben:

Definition 2.1 *Es sei* $\Psi_\xi \in \mathrm{Mat}(2L \times L,\, \mathbb{C})$ *und* $u \in \mathbb{C}^{2L}$. *Dann definiert man:*

$$u \in \Psi_\xi \Leftrightarrow \text{Es existiert ein } \widetilde{u} \in \mathbb{C}^L \text{ derart, dass } u = \Psi_\xi\, \widetilde{u} \text{ ist.}$$

Nun legt man die Gestalt von u_x genauer fest:

Definition 2.2 *Es sei* $y_x \in C^2\left(I, \mathbb{C}^L\right)$. p_x, q_x *und* v_x *seien in* $\mathrm{Mat}(L \times L,\, \mathbb{C})$ *dergestalt, dass* p_x, q_x *absolut stetig auf* I *sind, und* v_x *stetig auf* I *ist. Dann definiert man:*

$$u_x = \begin{pmatrix} y_x \\ p_x y_x' + q_x y_x \end{pmatrix} \qquad \text{für alle } x \in I.$$

Offensichtlich ist $u_x \in \mathbb{C}^{2L}$, *daher kann man Definition 2.1 anwenden.*

Noch eine letzte Definition ist nötig, bevor wir uns den Sturm-Liouville-Operatorn widmen können:

Definition 2.3 *Es sei* $\mathcal{J} = \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix}$ *die Weyl-Matrix (vgl. [deG97, Kap. 1, S. 20]).*

[1] Natürlich kann man alle Ergebnisse leicht auf den Fall $x \in [c_1,\, c_2],\ c_1,\ c_2 \in \mathbb{R}, c_1 < c_2$ übertragen.

Nun können wir die Randbedingungen beschreiben:

Definition 2.4 *Es gelte stets: $u_0 \in \Psi_{\xi_0}$ und $u_1 \in \Psi_{\xi_1}$. Weiterhin seien Ψ_{ξ_0} und Ψ_{ξ_1} so gewählt, dass sie die folgende Bedingung erfüllen:*

$$\Psi_{\xi_j}^* \mathcal{J} \Psi_{\xi_j} = 0 \qquad (j = 0,\ 1). \tag{2.1}$$

Im nächsten Kapitel wird sich zeigen, dass Ψ_{ξ_j} Repräsentanten einer Lagrange-Ebene *sind. Die so festgelegten Randbedingungen bezeichne ich mit $\mathcal{B}$ (für boundary conditions). Man sieht sofort, dass die Bedingung $\Psi_{\xi_j}^* \mathcal{J} \Psi_{\xi_j} = 0$ erfüllt ist, wählt man:*

$$\Psi_{\xi_j} = \begin{pmatrix} \xi_j \\ \mathbb{1}_L \end{pmatrix} \qquad \text{mit } \xi_j \in \operatorname{Her}(L, \mathbb{C}) \qquad (j = 0,\ 1). \tag{2.2}$$

Im Allgemeinen ist $\Psi_{\xi_0} \neq \Psi_{\xi_1}$, wählt man aber $\Psi_{\xi_0} = \Psi_{\xi_1} = \begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix}$ (offensichtlich erfüllt $\begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix}$ Bedingung (2.1)*), erhält man* Dirichlet-Bedingungen *am linken und am rechten Rand.*[2]

Hinweis: Im folgenden wird stets vorausgestzt, dass die Ebenen Ψ_{ξ_j} $(j = 0,\ 1)$ die Bedingung (2.1) erfüllen.

Aus den Definitionen 2.1 und 2.3 ergibt sich sofort, dass für alle $u,\ w \in \Psi_{\xi_j}$ $(j = 0,\ 1)$ gilt:

$$u^* \mathcal{J} w = \left(\Psi_{\xi_j} \widetilde{u}\right)^* \mathcal{J} \Psi_{\xi_j} \widetilde{w} = \widetilde{u}^* \underbrace{\Psi_{\xi_j}^* \mathcal{J} \Psi_{\xi_j}}_{=0} \widetilde{w} = 0. \tag{2.3}$$

Nun zur angkündigten Definition eines Sturm-Liouville-Operators:

Definition 2.5 (Sturm-Liouville-Operator)
Wir betrachten den Operator

$$\left(\widehat{\mathcal{L}} y\right)_x \equiv -\left(p_x y_x' + q_x y_x\right)' + q_x^* y_x' + v_x y_x, \qquad x \in I \tag{2.4}$$

[2] Ein Beispiel, das auf einen Sturm-Liouville-Operator mit Dirichlet-Randbedingungen führt, ist das Problem der schwingenden Saite, das qualitativ schon von Sokrates beschrieben wurde. Mathematisch geht es bei diesem Problem darum, die Werte $E \in \mathbb{R}$, für die die Differentialgleichung $-y'' = E\, y$ auf dem Intervall $I = [0, \pi]$ unter den Randbedingungen $y_0 = y_\pi = 0$ eine Lösung $y \neq 0$ hat, zu bestimmen. Die ersten Resultate zu diesem Problem gehen auf Daniel Bernoulli und Leonhard Euler zurück, siehe etwa [Wei87, Introduction, S. 8–11] oder [Sa99, Introduction, S. 1].

auf dem Definitionsbereich

$$\mathcal{D}\left(\widehat{\mathcal{L}}\right) = \left\{y \in C^2\left(I, \mathbb{C}^L\right) : \ u_0 \in \Psi_{\xi_0}, \ u_1 \in \Psi_{\xi_1}\right\} \subset L^2\left(I, \mathbb{C}^L\right).$$

Dabei sind p_x, q_x und $v_x \in \mathrm{Mat}(L \times L, \mathbb{C})$. p_x und q_x sind absolut stetig auf I,[3] *und v_x ist stetig auf I.*[4] *Außerdem gelten die folgenden Bedingungen:*

1. *p_x ist positiv definit.*[5]
2. *v_x ist hermitesch.*

In Abschnitt 2.3 wird gezeigt werden, dass $\widehat{\mathcal{L}}$ auf $\mathcal{D}\left(\widehat{\mathcal{L}}\right)$ wesentlich selbstadjungiert ist, und die eindeutige selbstadjungierte Erweiterung wird angegeben. Hier sei zunächst noch darauf hingewiesen, dass andere Differentialoperatoren, die auf eine gewöhnliche lineare Differentialgleichungen zweiter Ordnung führen, auf die Form von $\widehat{\mathcal{L}}$ gebracht werden können[6]. Um dies einzusehen, zeigt man zunächst:

Lemma 2.1 *Es sei $x \in I$. Jede Differentialgleichung der Form*

$$y_x'' + A_x^1 y_x' + A_x^0 y_x = B_x \tag{2.5}$$

(A_x^1 und A_x^0 sind hierbei Funktionen in $\mathrm{Mat}(L \times L, \mathbb{C})$) kann auf die Form

$$-\left(p_x y_x' + q_x y_x\right)' + q_x^* y_x' + v_x y_x = r_x \tag{2.6}$$

gebracht werden. Dabei sind p_x, q_x, v_x und r_x in $\mathrm{Mat}(L \times L, \mathbb{C})$. Außerdem sind p_x und q_x absolut stetig auf I, $p_x > 0$ für alle $x \in I$, und v_x sowie r_x sind stetig auf I.

Ebenso gilt die Umkehrung.

[3] Da p_x und q_x absolut stetig sind, existieren die Ableitungen p_x' und q_x' fast überall auf I.

[4] Im Allgemeinen reicht es zur Formulierung der Probleme sogar aus, v_x als lokal integrierbar über jedes endliche Intervall vorauszusetzen, (vgl. z. B. [CL55, Kap. 8.1, S. 208]). Hier wird v_x aber stetig gewählt. Zum einen ist es auf diese Weise einfacher, die Sätze aus der Theorie der linearen Differentialgleichungen anzuwenden. Wesentlich ist die Stetigkeit aber später für die Analyse des Verhaltens der Eigenphasen des Sturm-Liouville-Problems $\mathcal{L}y = Ey$ für $E \to -\infty$. Eine sehr genaue Analyse, in welchem Kontext welche Voraussetzungen nötig sind, um Selbstadjungiertheit zu garantieren, findet sich in [Wei05]. In Weidmanns Artikel werden nicht nur kompakte Intervalle I, sondern auch die Halbachsen und die komplette reelle Achse betrachtet.

[5] Im weiteren Verlauf der Arbeit wird positive Definitheit einer Matrix $T \in \mathrm{Mat}(L \times L, \mathbb{C})$ stets durch $T > 0$ erklärt. Es sei außerdem daran erinnert, dass sich aus $p_x > 0$ für alle $x \in I$ die folgenden Eigenschaften ergeben: Für $\alpha \in \mathbb{C}^L$, $\alpha \neq 0$ ist $\langle p_x \alpha | \alpha \rangle > 0$ für alle $x \in I$. Daraus ergibt sich zum einen die Selbstadjungiertheit von p_x (d. h. $p_x = p_x^*$ für alle $x \in I$) und zum anderen die Diagonalisierbarkeit von p_x mit Eigenwerten $\lambda_x^j > 0$ für alle $j = 1, \ldots, L$.

[6] Für das etwas einfachere Sturm-Liouville-Problem $\mathcal{L}^{CL} := \left(p_x y_x'\right)' + v_x y_x$ siehe dazu [CL55, Kap. 8.1, S. 208].

Beweis: Als erstes zeigt man, dass sich (2.6) auf die Form (2.5) bringen lässt. Die Gleichung $-\left(p_x y_x' + q_x y_x\right)' + q_x^* y_x' + v_x y_x = r_x$ bzw. $-p_x y_x'' - p_x' y_x' - q_x y_x' + q_x^* y_x' - q_x' y_x + v_x y_x = r_x$ lässt sich, da $p > 0$ ist, zu

$$y_x'' + \left(p_x^{-1} p_x' + p_x^{-1} q_x - p_x^{-1} q_x^*\right) y_x' + \left(p_x^{-1} q_x' y_x - p_x^{-1} v_x\right) y_x = -p_x^{-1} r_x$$

umformen. Wegen der Voraussetzungen existieren alle Koeffizientenmatrizen fast überall. Setzt man $A_x^1 = p_x^{-1} p_x' + p_x^{-1} q_x - p_x^{-1} q_x^*$, $A_x^0 = p_x^{-1} q_x' y_x - p_x^{-1} v_x$ und $B_x = -p_x^{-1} r_x$, so folgt die Gültigkeit von (2.6)⇒(2.5).

Für die Beweisrichtung (2.5)⇒(2.6) sei F_x eine Lösung der Differentialgleichung $F_x' = A_x^1$. Nun setze man $p_x = e^{F_x}$, $q_x = q_x^*$ beliebig, aber absolut stetig, $v_x = q_x' - e^{F_x} A_x^0$ und $r_x = e^{F_x} B_x$. Somit sind p_x, q_x, v_x und r_x absolut stetig, außerdem ist $p_x > 0$. Nun berechnet man:

$$\begin{aligned} & -\left(p_x y_x' + q_x y_x\right)' + q_x^* y_x' + v_x y_x - r_x \\ = \quad & -e^{F_x} y_x'' - e^{F_x} A_x^1 y_x' - q_x y_x' - q_x' y_x \\ & + q_x^* y_x' + \left(q_x' - e^{F_x} A_x^0\right) y_x - e^{F_x} B_x \\ \overset{q=q^*}{=} \quad & -e^{F_x}\left(y''_x + A_x^1 y_x' + A_x^0 y_x - B_x\right) = 0. \end{aligned}$$

Daraus folgt die Behauptung. □

Bemerkung 2.1 *Da sich bald herausstellen wird, dass* $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ *auf* $L^2\left(I, \mathbb{C}^L\right)$ *wesentlich selbstadjungiert ist (siehe Satz 2.1), genügt es zu zeigen, dass auf die Forderung* $v = v^*$ *nicht verzichtet werden kann, ohne die Symmetrie des Operators zu verletzen, um zu folgern, dass die eindeutige selbstadjungierte Erweiterung von* $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ *ein sehr allgemeiner*[7] *Operator ist, der alle obigen Forderungen bezüglich Glattheit und Selbstadjungiertheit erfüllt.*

Wichtige **Spezialfälle** des oben definierten Sturm-Liouville-Operators sind der *Schrödinger-Operator*[8]*:*

$$\mathcal{L}_S = y_x'' + v_x y_x$$

und der *Laplace-Operator:*

$$\mathcal{L}_L = y_x''.$$

[7] An den *allgemeinsten* Operator richtet sich nur die Forderung, dass $v_x \in L^1_{\mathrm{loc}}$ ist, die Stetigkeit muss man nicht zwingend voraussetzen.

[8] In der Quantenmechanik dient die Schrödinger-Gleichung dazu, das Verhalten von Quantenpartikeln mit Hilfe der spektralen Eigenschaften des Operators $\mathcal{L}_S$ zu bestimmen.

Häufig wählt man für einen Sturm-Liouville-Operator homogene Randbedingungen der Form $A_0^* y_0 + B_0^* y_0' = 0$ und $A_1 y_1 + B_1 y_1' = 0$, dabei sind A_0 und $B_0 \in \mathbb{C}^L$ (vgl. [deR05, S. 219]). Viele weitere Beispiele für Operatoren und ihre Randbedingungen finden sich in [Ev05a].

2.3 Eigenschaften des Sturm-Liouville-Operators

In diesem Abschnitt werden wir zeigen, dass $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ wesentlich selbstadjungiert auf $L^2\left(I, \mathbb{C}^L\right)$ ist und die selbstadjungierte Erweiterung $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ bestimmen. Anschließend werden wir uns mit der Beschränktheit von $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ auseinandersetzen. Doch vorweg müssen wir uns zunächst noch einmal mit den Randbedingungen beschäftigen:

Folgerung 2.1 *Erfüllen zwei Funktionen y, $\widetilde{y}$ die Randbedingungen $\mathcal{B}$, die gemäß Definition 2.4 gewählt sind, so folgt:*

$$\begin{aligned} & -\langle y_0 | p_0 \widetilde{y}_0' + q_0 \widetilde{y}_0 \rangle_{\mathbb{C}^L} + \langle p_0 y_0' + q_0 y_0 | \widetilde{y}_0 \rangle_{\mathbb{C}^L} = 0 \\ \text{und} \quad & -\langle y_1 | p_1 \widetilde{y}_1' + q_1 \widetilde{y}_1 \rangle_{\mathbb{C}^L} + \langle p_1 y_1' + q_1 y_1 | \widetilde{y}_1 \rangle_{\mathbb{C}^L} = 0 \end{aligned} \tag{2.7}$$

Denn: Nach Definition gilt:

$$u_0 = \begin{pmatrix} y_0 \\ p_0 y_0' + q_0 y_0 \end{pmatrix} \in \Psi_{\xi_0} \quad \text{und} \quad \widetilde{u}_0 = \begin{pmatrix} \widetilde{y}_0 \\ p_0 \widetilde{y}_0' + q_0 \widetilde{y}_0 \end{pmatrix} \in \Psi_{\xi_0}.$$

Damit erhält man:

$$\begin{aligned} 0 &\overset{(2.3)}{=} u_0^* \mathcal{J} \widetilde{u}_0 \\ &= \begin{pmatrix} y_0 \\ p_0 y_0' + q_0 y_0 \end{pmatrix}^* \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix} \begin{pmatrix} \widetilde{y}_0 \\ p_0 \widetilde{y}_0' + q_0 \widetilde{y}_0 \end{pmatrix} \\ &= -y_0^* \left(p_0 \widetilde{y}_0' + q_0 \widetilde{y}_0\right) + \left(p_0 y_0' + q_0 y_0\right)^* \widetilde{y}_0 \\ &= -\langle y_0 | p_0 \widetilde{y}_0' + q_0 \widetilde{y}_0 \rangle + \langle p_0 y_0' + q_0 y_0 | \widetilde{y}_0 \rangle. \end{aligned}$$

Die Behauptung für $x = 1$ ergibt sich analog. □

2.3.1 Selbstadjungiertheit

Satz 2.1 *Es gilt:*

$$\widehat{\mathcal{L}} \text{ ist } \textit{wesentlich selbstadjungiert} \text{ auf } \mathcal{D}\left(\widehat{\mathcal{L}}\right) \subset L^2\left(I, \mathbb{C}^L\right).$$

Beweis: Um die wesentliche Selbstadjungiertheit zu beweisen, genügt es zu zeigen, dass der Operator symmetrisch auf $\mathcal{D}\left(\widehat{L}\right)$ ist, dass $\mathcal{D}\left(\widehat{L}\right)$ dicht in $L^2\left(I, \mathbb{C}^L\right)$ ist und dass $\ker\left(\widehat{\mathcal{L}} \pm i\right) = \{0\}$ ist, vgl. [Wer04, Kap. VII., S. 343].

1. **$\widehat{\mathcal{L}}$ ist symmetrisch auf $\mathcal{D}\left(\widehat{L}\right)$:** Durch partielle Integration ergibt sich (dabei beachte man die Randbedingungen und die Tatsache, dass $v = v^*$ ist):

$$\begin{aligned}
& \langle \mathcal{L}y|\widetilde{y}\rangle_{L^2(I,\mathbb{C}^L)} - \langle y|\mathcal{L}\widetilde{y}\rangle_{L^2(I,\mathbb{C}^L)} \\
= & \int_I \mathrm{dx}\Big(\langle -\big(p_x y_x' + q_x y_x\big)' + q_x^* y_x' + v_x y_x|\widetilde{y}_x\rangle_{\mathbb{C}^L} \\
& \qquad -\langle y_x| - \big(p_x \widetilde{y}_x + q_x \widetilde{y}_x\big)' + q_x^* \widetilde{y}_x{}' + v_x \widetilde{y}_x\rangle_{\mathbb{C}^L}\Big) \\
\overset{v=v^*}{=} & \int_I \mathrm{dx}\Big(\langle -\big(p_x y_x' + q_x y_x\big)'|\widetilde{y}_x\rangle - \langle y_x| - \big(p_x \widetilde{y}_x' + q_x \widetilde{y}_x\big)'\rangle\Big) \\
& + \int_I \mathrm{dx}\Big(\langle q_x^* y_x'|\widetilde{y}_x\rangle - \langle y_x|q_x^* \widetilde{y}_x'\rangle\Big) \\
\overset{\text{part. Int., (2.7)}}{=} & \int_I \mathrm{dx}\langle p_x y_x' + q_x y_x|\widetilde{y}_x{}'\rangle - \int_I \mathrm{dx}\langle y_x'|p_x \widetilde{y}_x{}' + q_x \widetilde{y}_x\rangle \\
& + \int_I \mathrm{dx}\langle q_x^* y_x'|\widetilde{y}_x\rangle - \int_I \mathrm{dx}\ \langle y_x|q_x^* \widetilde{y}_x'\rangle \\
= & \int_I \mathrm{dx}\ \langle p_x y_x'|\widetilde{y}_x'\rangle - \int_I \mathrm{dx}\ \langle y_x'|p_x \widetilde{y}_x'\rangle + \int_I \mathrm{dx}\ \langle q_x y_x|\widetilde{y}_x'\rangle \\
& - \int_I \mathrm{dx}\ \langle y_x'|q_x \widetilde{y}_x\rangle + \int_I \mathrm{dx}\ \langle q_x^* y_x'|\widetilde{y}_x\rangle - \int_I \mathrm{dx}\ \langle y_x|q_x^* \widetilde{y}_x'\rangle \overset{p=p^*}{=} 0
\end{aligned}$$

Somit ist der Operator $\widehat{\mathcal{L}}$ symmetrisch.

2. **$\mathcal{D}\left(\widehat{L}\right)$ ist dicht in $L^2\left(I, \mathbb{C}^L\right)$:** Diese Behauptung ergibt sich sofort daraus, dass $C^2_K\left((0,\ 1), \mathbb{C}^L\right) \subset \mathcal{D}\left(\widehat{\mathcal{L}}\right)$ ist. Da $C^2_K\left((0,\ 1), \mathbb{C}^L\right)$ in $L^2\left(I, \mathbb{C}^L\right)$ dicht ist, folgt die Aussage.

3. **$\ker\left(\widehat{\mathcal{L}}^* \pm i\mathbb{1}\right) = \{0\}$:** Der Beweis verläuft analog zu [Pe88, Kap. 6, S. 250].

Daher ist $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ auf $L^2\left(I, \mathbb{C}^L\right)$ wesentlich selbstadjungiert. □

Bemerkung 2.2 *Nun nutzt man aus, dass ein Operator genau dann wesentlich selbstadjungiert ist, wenn sein Abschluss selbstadjungiert ist. Um den selbstadjungierten Operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ zu bestimmen, muss man also den Abschluss von $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ kennen. Weiterhin sei darauf hingewiesen, dass $\left(\widehat{\mathcal{L}}^{**}, \mathcal{D}\left(\widehat{\mathcal{L}}^{**}\right)\right)$ die kleinste abgeschlossene Erweiterung von $\widehat{\mathcal{L}}$ ist, da $\left(\widehat{\mathcal{L}}^{*}, \mathcal{D}\left(\widehat{\mathcal{L}}^{*}\right)\right)$ wegen der Symmetrie dicht definiert ist, vgl. [Wer04, Kap. VII, S. 339]. Wegen der wesentlichen Selbstadjungiertheit gibt es aber nur genau eine selbstadjungierte Erweiterung von $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$, und diese ist nach obigen Vorbemerkungen $\left(\widehat{\mathcal{L}}^{**}, \mathcal{D}\left(\widehat{\mathcal{L}}^{**}\right)\right)$.*

Satz 2.2 *Der Operator*

$$(\mathcal{L}y)_x \equiv -\left(p_x y'_x + q_x y_x\right)' + q^*_x y'_x + v_x y_x, \qquad x \in I$$

auf dem Definitionsbereich

$$\mathcal{D}\left(\mathcal{L}\right) = \left\{y \in C^1\left(I, \mathbb{C}^L\right) : y' \in AC\left(I, \mathbb{C}^L\right),\ u_j \in \Psi_{\xi_j},\ (j = 0,\ 1)\right\}$$

$\left(\mathcal{D}\left(\mathcal{L}\right) \subset L^2\left(I, \mathbb{C}^L\right)\right)$ *ist selbstadjungiert.*

Beweis: Vorweg sei festgehalten, dass $(\mathcal{L}, \mathcal{D})$ eine Erweiterung von $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ ist.

Als erstes demonstriert man, dass $(\mathcal{L}, \mathcal{D})$ *abgeschlossen* ist. Wie in [Pe88, Kap. 6, S. 252] kann man zeigen, dass für jede Folge $(y_n)_{n \in \mathbb{N}} \in L^2(I, \mathbb{C}^L)$, die stark gegen ein $y \in L^2(I, \mathbb{C}^L)$ konvergiert und die die Bedingung $\mathcal{L}y_n \overset{n\to\infty}{\to} \widetilde{y}$ stark ($\widetilde{y} \in L^2(I, \mathbb{C}^L)$) erfüllt, gilt:

$$\mathcal{L}y_n \to \mathcal{L}y \text{ stark} \qquad (n \to \infty).$$

Daher ist $(\mathcal{L}, \mathcal{D})$ ein abgeschlossener Operator.

Nun zeigt man, dass $\mathcal{D}\left(\widehat{\mathcal{L}}^{**}\right) = \mathcal{D}(\mathcal{L})$ gilt: Man kann beweisen, dass für $z \in \mathbb{C}$ und $y \in \mathcal{D}\left(\widehat{\mathcal{L}}^{*}\right)$ gilt:[9]

$$\left(\widehat{\mathcal{L}}^{*} - z\mathbb{1}\right) y = 0 \quad \Leftrightarrow \quad (\mathcal{L}^{*} - z\mathbb{1})\, y = 0.$$

Die Beweisidee dazu findet man wiederum in [Pe88, Kap. 6, S. 252/253]. Das bedeutet, die Lösungsmengen von $\left(\widehat{\mathcal{L}}^{*} - z\mathbb{1}\right) y = 0$ und $(\mathcal{L}^{*} - z\mathbb{1})\, y = 0$ sind identisch. Daher sind auch die Bilder von $\mathcal{L} - \overline{z}\,\mathbb{1}$ und $\widehat{\mathcal{L}}^{**} - \overline{z}\,\mathbb{1}$ identisch, da beide jeweils orthogonale Unterräume zu den obigen Lösungsmengen sind. Somit gilt:

$$\mathcal{D}\left(\widehat{\mathcal{L}}^{**}\right) = \mathcal{D}(\mathcal{L}),$$

d. h. $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ ist eine abgeschlossene Erweiterung von $\left(\widehat{\mathcal{L}}, \mathcal{D}\left(\widehat{\mathcal{L}}\right)\right)$ und wegen Bemerkung 2.2 folgt die Behauptung. □

Folgerung 2.2 *Aus dem Beweis zu den Satz 2.1 (Symmetrie) und aus Satz 2.2 (Gestalt des Operators) wird unmittelbar ersichtlich, dass auf die Forderung $v = v^*$ nicht verzichtet werden kann, da sonst der Operator $(\mathcal{L}, \mathcal{D})$ im Allgemeinen nicht selbstadjungiert ist.*

Folgerung 2.3 *Wegen der Selbstadjungiertheit von $\mathcal{L}$ gilt:*

- *Das Spektrum von $(\mathcal{L}, \mathcal{D})$, bezeichnet mit $\sigma(\mathcal{L})$, erfüllt:*

$$\sigma(\mathcal{L}) \subset \mathbb{R}. \tag{2.8}$$

  Ein Beweis dazu findet sich beispielsweise in [Wer04, Kap. VII, S. 349].

- *Eigenvektoren, die zu unterschiedlichen Eigenwerten gehören, stehen orthogonal aufeinander. Einen Beweis dazu gibt [HS96, Kap. 13, S. 133/134].*

2.3.2 Beschränktheit

Der folgende Satz zeigt, dass der Sturm-Liouville-Operator $(\mathcal{L}, \mathcal{D})$ unter bestimmten Voraussetzungen nach unten beschränkt ist.

[9] Man beachte dabei, dass auch $\mathcal{L}^* y$ somit wohldefiniert ist.

Satz 2.3 *Für alle $x \in I$ sei $q_x = q_x^*$, und es gelte $q_x \in C^1(I, \mathrm{Mat}(L \times L, \mathbb{C}))$. Setzt man weiterhin Dirichlet-Randbedingungen am linken und rechten Rand voraus, so existiert ein $\alpha \in \mathbb{R}$ für das gilt:*

$$\mathcal{L} \geq \alpha.$$

Das bedeutet, für alle $\phi \in \mathcal{D}(\mathcal{L})$ gilt:

$$0 \leq \langle \phi | (\mathcal{L} - \alpha \mathbb{1}) | \phi \rangle.$$

Beweis: Es sei $\phi_x \in \mathcal{D}(\mathcal{L})$. Mit partieller Integration erhält man (analog zum Beweis der Symmetrie in Satz 2.1):

$$\begin{aligned} &\langle \phi_x | \mathcal{L} \phi_x \rangle \\ &= \underbrace{-\langle \phi_x | p_x \phi_x' + q_x \phi_x \rangle \Big|_0^1}_{=:C_1} + \underbrace{\int\limits_I \mathrm{dx} \left(\langle \phi_x' | p_x \phi_x' \rangle + \langle \phi_x | v_x \phi_x \rangle \right)}_{=:C_2} \\ &\quad + \underbrace{\int\limits_I \mathrm{dx} \left(\langle \phi_x' | q_x \phi_x \rangle + \langle \phi_x | q_x^* \phi_x' \rangle \right)}_{=:C_3} \end{aligned} \tag{2.9}$$

Nun kann man jedes dieser Glieder C_i $(i = 1,\ 2,\ 3)$ nach unten abschätzen:

$\boldsymbol{C_1} = 0$, denn unter Dirichlet-Randbedingungen gilt: $\phi_0 = \phi_1 = 0$.

Die Abschätzung für $\boldsymbol{C_2}$ ergibt sich dadurch, dass $p_x > 0$ ist für alle $x \in I$. Daher ist $\langle \phi' | p_x \phi' \rangle > 0$ und ebenso ist $\int_I \mathrm{dx} \langle \phi_x' | p_x \phi_x' \rangle > 0$. Außerdem beachte man, dass v stetig und I kompakt ist, damit existiert $\min\limits_{x \in I} v_x =: \alpha_2$. Somit gilt:

$$C_2 \geq \int\limits_I \mathrm{dx}\ \langle v_x \phi_x | \phi_x \rangle \geq \alpha_2 \int\limits_I \mathrm{dx}\ \langle \phi_x | \phi_x \rangle.$$

Insgesamt erhält man: $C_2 \geq \alpha_2$.

Um $\boldsymbol{C_3}$ abzuschätzen benötigt man die Voraussetzungen an q. Da $q = q^*$ ist, gilt:

$$C_3 = \int\limits_I \mathrm{dx} \left(\langle \phi_x' | q_x \phi_x \rangle + \langle \phi_x | q_x \phi_x' \rangle \right).$$

Mit partieller Integration sieht man, dass folgende Formel gilt:

$$\langle \phi_x | q \phi_x \rangle \Big|_0^1 = \int\limits_I \mathrm{dx} \left(\langle \phi_x | q_x \phi_x \rangle \right)' = \int\limits_I \mathrm{dx} \left(\langle \phi_x' | q_x \phi_x \rangle + \langle \phi_x | q_x \phi_x' \rangle + \langle \phi_x | q_x' \phi_x \rangle \right).$$

Daraus folgt:

$$\int_I \mathrm{dx}\left(\langle\phi'_x|q_x\phi_x\rangle + \langle\phi_x|q_x\phi'_x\rangle\right) = \underbrace{\langle\phi_x|q_x\phi_x\rangle\Big|_0^1}_{=:C_4} - \underbrace{\int_I \mathrm{dx}\langle\phi_x|q'_x\phi_x\rangle}_{=:C_5} .$$

Analog zu obiger Argumentation für C_1 sieht man, dass wegen der Dirichlet-Bedingungen $C_4 = 0$ ist. Um die Existenz einer unteren Schranke für C_5 zu beweisen, benötigt man die zweite Voraussetzung an q: Da q in $C^1([0,\ 1], \mathrm{Mat}(L \times L,\ \mathbb{C}))$ ist, ist q'_x für alle $x \in I$ beschränkt, und damit ist C_5 nach unten beschränkt. Zusammengefasst ergibt sich die Existenz eines $\alpha_3 \in \mathbb{R}$, für das gilt: $C_3 \geq \alpha_3$.

Setzt man diese Abschätzungen wieder in (2.9) ein, ergibt sich:

$$\langle\phi|\mathcal{L}\phi\rangle \geq \alpha,$$

wobei $\alpha = \alpha_2 + \alpha_3$ ist, und man hat, wie behauptet eine untere Schranke gefunden, denn es gilt:

$$\langle\phi|\mathcal{L} - \alpha\mathbb{1}|\phi\rangle \geq 0.$$

Der Satz ist somit bewiesen. □

Definition 2.6 (Das reguläre Sturm-Liouville-Problem)
Sei $E \in \mathbb{R}$ und $y \in \mathcal{D}(\mathcal{L})$, wobei $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ wie in Satz 2.2 definiert ist. Die Lösung des Problems

$$\mathcal{L}y = Ey \tag{2.10}$$

bezeichnet man als Lösung des ***Sturm-Liouville-Problems.*** *Da nach Satz 2.2 der Operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ selbstadjungiert ist, genügt es, $E \in \mathbb{R}$ zu betrachten, für $E \in \mathbb{C} \setminus \mathbb{R}$ hat das Problem nach* (2.8) *nur die triviale Lösung.*

2.4 Das Problem als System erster Ordnung

Wie so oft bei Differentialgleichungen höherer Ordnung, bietet es sich auch für die durch den Sturm-Liouville-Operator induzierte Differentialgleichung zweiter Ordnung an, das Problem mittels der nun folgenden Definition auf ein äquivalentes System erster Ordnung zu bringen (vgl. [Bo56]). Damit erhält man ein System der Form $u'_x = \mathfrak{t}^E_x u_x$, wobei sowohl u_x als auch u'_x in $\mathbb{C}^{2L}$ sind.

Definition 2.7 *Sei $y_x \in \mathcal{D}(\mathcal{L})$. Man definiere u_x als Funktion $I \to \mathbb{C}^{2L}$ durch:*

$$u_x = \begin{pmatrix} y_x \\ p_x y'_x + q_x y_x \end{pmatrix} . \tag{2.11}$$

Lemma 2.2 *Das Sturm-Liouville-Problem $\mathcal{L}y = Ey$ ist äquivalent zu*

$$u'_x = \mathfrak{t}_x^E \, u_x. \tag{2.12}$$

Dabei ist

$$\mathfrak{t}_x^E = \begin{pmatrix} -p_x^{-1} q_x & p_x^{-1} \\ v_x - q_x^* p_x^{-1} q_x - E \mathbb{1}_L & q_x^* p_x^{-1} \end{pmatrix}. \tag{2.13}$$

Hinweis: Um die Notation übersichtlicher zu gestalten, werde ich im folgenden die Abhängigkeit aller hier behandelten Funktionen von x, nicht mehr explizit ausschreiben – es sei denn, diese Abhängigkeit ist von entscheidender Bedeutung. Dennoch besteht diese Abhängigkeit weiterhin überall!

Nun zum **Beweis** des Lemmas. Es gilt:

$$u' = \begin{pmatrix} y' \\ (py' + qy)' \end{pmatrix}$$

Andererseits ist

$$\begin{aligned} \mathfrak{t}^E u &= \begin{pmatrix} -p^{-1} q & p^{-1} \\ v - q^* p^{-1} q - E\mathbb{1} & q^* p^{-1} \end{pmatrix} \begin{pmatrix} y \\ py' + qy \end{pmatrix} \\ &= \begin{pmatrix} -p^{-1} qy + p^{-1}(py' + qy) \\ (v - q^* p^{-1} q - E)y + q^* p^{-1}(py' + qy) \end{pmatrix} \\ &= \begin{pmatrix} -p^{-1} qy + y' + p^{-1} qy \\ vy - q^* p^{-1} qy - Ey + q^* y' + q^* p^{-1} qy \end{pmatrix} \\ &= \begin{pmatrix} y' \\ vy - Ey + q^* y' \end{pmatrix} \end{aligned}$$

Daraus folgt:

$$\begin{aligned} u' &= \mathfrak{t}^E u \\ \Leftrightarrow \quad (py' + qy)' &= vy - Ey + q^* y' \\ \Leftrightarrow \quad -(py' + qy)' + q^* y' + vy &= Ey \\ \Leftrightarrow \quad \mathcal{L}y &= Ey. \end{aligned}$$

Also entspricht jede Lösung von (2.12) auch einer Lösung von $\mathcal{L}y = Ey$. □

Hinweis: Im folgenden betrachte man stets das System

$$u'_x = \mathfrak{t}_x^E u_x,$$

dabei ist u_x definiert wie in (2.11) und $\mathfrak{t}_x^E$ wie in (2.13). Die Ebene, die von den Lösungen des Problems $u'_x = \mathfrak{t}_x^E u_x$, die die Randbedingung am linken Rand erfüllen, für ein festes $x \in I$ aufgespannt wird, bezeichne ich mit Φ_x^E (d. h. für alle

$u_x \in \Phi_x^E$ gilt $u_x' = \mathfrak{t}_x^E u_x$, und die Anfangsbedingung $u_0 \in \Psi_{\xi_0}$ mit $\Psi_{\xi_0}^* \mathcal{J} \Psi_{\xi_0} = 0$ ist erfüllt$\big)$. Speziell ist $\Phi_0^E = \Psi_{\xi_0}$. Im dritten Kapitel werden wir näher auf die Eigenschaften dieser Lösungsebenen eingehen. Es wird sich herausstellen, dass diese Ebenen *Lagrange-Ebenen* sind. Sei weiterhin Σ^E die Anzahl der linear unabhängigen Lösungen von $\mathcal{L}y = Ey$ unter beiden Randbedingungen $\mathcal{B}$. Dann bilden die Punkte E, an denen gilt: $\Sigma^E \neq 0$ das *Spektrum* von (2.10).

2.5 Analytizität des Systems erster Ordnung in *E*

Im letzten Abschnitt dieses Kapitels steht die Analytizität von (2.10) für den Parameter E im Vordergrund.[10] Dazu betrachte man das System erster Ordnung $u_x' = \mathfrak{t}_x^E u_x$. Um die Analytizität in E zu beweisen, greift man auf einige Existenz- und Eindeutigkeitsaussagen aus der Theorie der linearen Differentialgleichungen zurück, die sich in den gängigen Lehrbüchern finden, siehe beispielsweise [CL55, Kap. 2] oder [Wa93, Kap. 2]. Die für diese Arbeit gewählten Formulierungen sind vor allem [Mu03, Kap. 23] entnommen. Alle verwendeten Aussagen werden in Anhang C vorgestellt.

Satz 2.4 *Das Sturm-Liouville-Problem – formuliert als System erster Ordnung gemäß* (2.12) *– hat auf dem Intervall I eine eindeutige Lösung, die analytisch in E ist.*

Beweis: Es genügt nach dem Existenz- und Eindeutigkeitssatz von Picard-Lindelöf und Satz C.3 zu zeigen, dass $u_x' = \mathfrak{t}_x^E u_x$ einer lokalen Lipschitz-Bedingung genügt und analytisch in E ist. Zu diesem Zweck betrachte man die Matrix

$$\mathfrak{t}_x^E = \begin{pmatrix} -p_x^{-1} q_x & p_x^{-1} \\ v_x - q_x^* p_x^{-1} q_x - E\mathbb{1} & q_x^* p_x^{-1} \end{pmatrix}$$

etwas genauer. Nach Definition des Sturm-Liouville-Problems (2.5) sind p_x, q_x und v_x mindestens stetige Funktionen in $\mathrm{Mat}(L \times L, \mathbb{C})$. Damit sind auch p_x^{-1} und q_x^* stetig. Daher ist $f(x, u_x) = \mathfrak{t}_x^E u_x$ in allen Komponenten stetig differenzierbar. Mit Bemerkung C.1 folgt, dass die Lipschitz-Bedingung erfüllt ist. Nach dem Satz von Picard-Lindelöf (Satz C.2) ist damit die Existenz und Eindeutigkeit einer Lösung sichergestellt.

[10] Die Analytizität der Lösungen in x wird beispielsweise in [Wei87, Kap. 2, S. 23–25] thematisiert, doch für diese Arbeit ist nur die Analytizität in E wesentlich.

Man sieht sofort, dass $\mathfrak{t}_x^E u_x$ darüber hinaus analytisch in E ist. Also ist nach Satz C.3 jede Lösung des Sturm-Liouville-Problems analytisch in E. $\square$

Damit sind die Grundzüge des Sturm-Liouville-Problems umrissen. Im nächsten Kapitel geht es darum, die von den Lösungen aufgespannten Ebenen zu charakterisieren, um in den dann folgenden Kapiteln Aussagen über das Verhalten der Eigenwerte zu machen.

Kapitel 3

Hermitesch symplektische Geometrie

Structures are the weapons of the mathematician.

Nicholas Bourbaki, [Ha57].

Dieses Kapitel ist der hermitesch symplektischen Geometrie gewidmet. Zunächst wird eine kurze Einführung gegeben, eine grundlegendere Diskussion findet sich in Anhang A. Darüber hinaus werde ich die Struktur der Grassmann-Mannigfaltigkeit und der Lie-Algebra erläutern, um das Bild zu komplettieren. Es wird sich zeigen, dass diese Strukturen den geeigneten Rahmen bilden, um das Verhalten der Lösungen zu beschreiben und das Oszillationstheorem herzuleiten.

In der Literatur wird meist die reelle symplektische Geometrie betrachtet, während in dieser Arbeit die hermitesch symplektische Geometrie, d. h. eine Geometrie über *komplexe* Vektorräume, herangezogen wird. In Anhang B wird gezeigt, dass sowohl die reelle als auch die quaternionisch symplektische Geometrie durch Forderung zusätzlicher Symmetrieeigenschaften aus der hermitesch symplektischen Geometrie hervorgehen. Diese Ergebnisse gehen auf [SB06] zurück. Im letzten Abschnitt dieses Kapitels werden die Strukturen der hermitesch symplektischen Geometrie auf das Sturm-Liouville-System übertragen.

3.1 Die hermitesch symplektische Standardform

Im ersten Abschnitt dieses Kapitels wird die hermitesch symplektische Geometrie erörtert. Diese Geometrie wird im Allgemeinen über eine Transformationsgruppe definiert, die hier als Standardform bezeichnet wird. In Anhang A wird gezeigt, dass es genügt, sich nur mit dieser hermitesch symplektischen Standardform zu beschäftigen, da alle anderen symplektischen Formen zu ihr äquivalent sind.

Definition 3.1 *Es seien* $\nu_1,\ \nu_2 \in \mathbb{C}^{2L}$*. Die* hermitesch symplektische Standardform *ist definiert als:*

$$\omega_0(\nu_1,\nu_2) = \nu_1^* \mathcal{J} \nu_2 \qquad \text{mit} \qquad \mathcal{J} = \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix}. \tag{3.1}$$

Eine Matrix $\mathfrak{M}$ *erhält die Form* (3.1) *also genau dann, wenn gilt:*

$$\mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J}.$$

Somit ist die hermitesch symplektische Standardgruppe $\mathfrak{SP}(2L,\mathbb{C})$ *gegeben durch:*

$$\mathfrak{SP}(2L,\mathbb{C}) = \{\mathfrak{M} \in \mathrm{Mat}(2L \times 2L\,,\mathbb{C}) : \ \mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J}\}\,. \tag{3.2}$$

Lemma 3.1 *Die Definition der hermitesch symplektischen Standardgruppe führt zu folgender Charakterisierung:*

$$\mathfrak{SP}(2L,\mathbb{C}) = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Mat}(2L \times 2L,\ \mathbb{C}) : \right. \\ \left. A^*C = C^*A,\ A^*D - C^*B = \mathbb{1},\ B^*D = D^*B \right\} \tag{3.3}$$

Für den **Beweis** schreibe man $\mathfrak{SP}(2L,\mathbb{C}) \ni \mathfrak{M}$ in Blockmatrixform, d. h. $\mathfrak{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Damit gilt:

$$\begin{aligned} \begin{pmatrix} A & B \\ C & D \end{pmatrix}^* \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \\ \Leftrightarrow \quad \begin{pmatrix} -A^*C + C^*A & -A^*D + C^*B \\ -(-D^*A + B^*C) & -B^*D + D^*B \end{pmatrix} &= \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \end{aligned}$$

Wenn man beachtet, dass $\mathbb{1}^* = \mathbb{1}$ ist, folgt die Behauptung. □

3.1.1 Folgerungen

- Man sieht sofort, dass $\mathcal{J}^2 = -\mathbb{1}_{2L}$ und $\mathcal{J}^* = -\mathcal{J}$ ist. Außerdem folgt aus Lemma 3.3, dass die Determinante hermitesch symplektischer Matrizen gleich 1 ist.

- Die Inverse zu einer Matrix $\mathfrak{M} \in \mathfrak{SP}(2L, \mathbb{C})$ ist gegeben durch

$$\mathfrak{M}^{-1} = \begin{pmatrix} D^* & -B^* \\ -C^* & A^* \end{pmatrix}. \tag{3.4}$$

  Denn:

$$\begin{aligned} \mathfrak{M}\mathfrak{M}^{-1} &= \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D^* & -B^* \\ -C^* & A^* \end{pmatrix} \\ &= \begin{pmatrix} AD^* - BC^* & -AB^* + BA^* \\ CD^* - C^*D & -CB^* + DA^* \end{pmatrix}. \end{aligned}$$

  Mit (3.3) folgt:

$$\begin{aligned} AD^* - BC^* &= \mathbb{1}_L \\ -CB^* + DA^* = (AD^* - BC^*)^* &= (\mathbb{1}_L)^* = \mathbb{1}_L \\ CD^* - C^*D &= 0_L \\ -AB^* + BA^* &= 0_L. \end{aligned}$$

  Also ist $\mathfrak{M}\mathfrak{M}^{-1} = \mathbb{1}_{2L}$. Analog rechnet man nach, dass $\mathfrak{M}^{-1}\mathfrak{M} = \mathbb{1}_{2L}$ ist.

- Hat man eine hermitesch symplektische Basis von $\left(\mathbb{C}^{2L}, \omega\right)$ ausgewählt, so kann man eine lineare Abbildung $m : \mathbb{C}^{2L} \to \mathbb{C}^{2L}$ mit einer Matrix $\mathfrak{M}$ in dieser Basis identifizieren.

- Mit Hilfe von (3.1) lassen sich die bereits im Anhang definierten *Lagrange-Ebenen* konkretisieren. Nach Definition A.3 sind Lagrange-Ebenen Unterräume eines Vektorraums, die sowohl isotrop als auch coisotrop sind. Man betrachte einen hermitesch symplektischen Vektorraum $(\mathcal{V}, \omega)$ über $\mathbb{C}$ mit $\dim(\mathcal{V}) = 2L$. Mittels der gewählten Konventionen ist die Isomorphie von $\mathfrak{SP}(\mathcal{V}, \omega)$ und $\mathfrak{SP}(2L, \mathbb{C})$ sichergestellt, und es gilt (vgl. [FJN03]): Φ ist genau dann eine Lagrange-Ebene, wenn $\Phi \subset \mathbb{C}^{2L \times L}$ mit $\nu_1^* \mathcal{J} \nu_2 = 0$ für alle ν_1, $\nu_2 \in \Phi$. Zusammengefasst ergibt sich: $\Phi^* \mathcal{J} \Phi = 0$.

- Man kann zeigen, dass $\mathfrak{SP}(2L, \mathbb{C})$ zusammenhängend ist (siehe [deG97, Kap. 1, S. 20/21]).

Lemma 3.2 *Die Gruppe $\mathfrak{SP}(2L,\mathbb{C})$ ist stabil unter Bildung der Adjungierten.*

Beweis (vgl. [deG97, Kap. 1]): Nach (3.2) ist $\mathfrak{M} \in \mathfrak{SP}(2L,\mathbb{C}) \Leftrightarrow \mathfrak{M}^*\mathcal{J}\mathfrak{M} = \mathcal{J}$. Da $\mathfrak{SP}(2L,\mathbb{C})$ eine Gruppe ist, ist auch $\mathfrak{M}^{-1} \in \mathfrak{SP}(2L,\mathbb{C})$, d. h. $(\mathfrak{M}^{-1})^*\mathcal{J}\mathfrak{M}^{-1} = \mathcal{J}$. Wichtig ist, dass $\dim(\mathcal{J}) = 2L$ ist, d. h. beide Seiten der Gleichung sind invertierbar. Invertiert man beide Seiten der Gleichung, erhält man $\mathfrak{M}\mathcal{J}^{-1}\mathfrak{M}^* = \mathcal{J}^{-1}$ bzw. $\mathfrak{M}\mathcal{J}\mathfrak{M}^* = \mathcal{J}$. □

3.2 Die Grassmann-Mannigfaltigkeit

Lagrange-Ebenen im hermitesch symplektischen Raum

In diesem Abschnitt geht es um die schon kurz vorgestellten *Lagrange-Ebenen.* Dieser Teil der Arbeit stellt einen Rahmen vor, der es später erlauben wird, die Dynamik der Lösungen des Sturm-Liouville-Problems zu beschreiben, denn es wird sich zeigen, dass die Lösungsebenen des Sturm-Liouville-Problems an einem Punkt $x \in I$ tatsächlich Lagrange-Ebenen sind.

Nach Definition A.3 ist ein Unterraum Φ eines Vektorraums $(\mathcal{V},\omega)$ genau dann eine Lagrange-Ebene, wenn Φ sowohl isotrop als auch coisotrop ist. Aufgrund der Isotropie verschwindet ω identisch auf Φ. Aus der Dimensionsformel (A.1) ergibt sich außerdem, dass $\dim(\Phi) = \frac{1}{2}\dim(\mathcal{V})$ ist.

Zunächst betrachten wir die L-dimensionalen Unterräume eines komplexen Vektorraums $\mathbb{C}^{2L}$. Wählt man Basisvektoren aus, die einen solchen Unterraum aufspannen, bilden die Basisvektoren eine $2L\times L$-Matrix, die ich mit Φ bezeichne – in Anbetracht des vorher Bemerkten recht suggestiv. Offensichtlich ist $\mathrm{Rang}\,(\Phi) = L$. Darüber hinaus ist die Ebene natürlich unabhängig von der konkreten Wahl der Basis (und somit von der explizit gewählten Form von Φ).

Definition 3.2 *Um die Grassmann-Mannigfaltigkeit der Lagrange-Ebenen zu definieren, wird folgende Äquivalenzrelation eingeführt:*

$$\Phi \sim \Psi \;\Leftrightarrow\; \text{Es existiert ein } c \in Gl(L,\mathbb{C}) \quad \text{mit} \quad \Phi = \Psi\, c. \tag{3.5}$$

Man beachte, dass diese Definition in Analogie zu Definition 2.1 steht, in der das Vorgehen bei der Wahl von Ebenen als Randbedingungen definiert wurde.

Definition 3.3 *Die Grassmann-Mannigfaltigkeit der Lagrange-Ebenen* $\mathbb{G}_L^{\mathbb{C}}$ *ist definiert als die Menge der Äquivalenzklassen* $\sim$ *aus* (3.5)*:*

$$\mathbb{G}_L^{\mathbb{C}} = \{[\Phi]_\sim : \; \Phi \in \mathrm{Mat}(2L \times L,\, \mathbb{C}),\; \mathrm{Rang}(\Phi) = L\}\,.$$

Als hermitesch symplektische Struktur auf dieser Mannigfaltigkeit wählt man die hermitesch symplektische Standardform gemäß Definition 3.1.

Lemma 3.3 $[\Phi]_\sim \in \mathbb{G}_L^{\mathbb{C}}$ *ist genau dann eine Lagrange-Ebene, wenn gilt:*

$$\Phi^* \mathcal{J} \Phi = 0 \qquad \text{für alle } \Phi \in [\Phi]_\sim.$$

Weitere Folgerungen erhält man, indem man Φ mit Hilfe von zwei Matrizen a und b schreibt:

Folgerung 3.1 *Es sei* $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ *mit* $a,\, b \in \mathrm{Mat}(L \times L,\, \mathbb{C})$*. Die Bedingung* $\Phi^* \mathcal{J} \Phi = 0$ *wird in dieser Notation zu:*

$$0 = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = -a^* b + b^* a. \tag{3.6}$$

Äquivalent zu $\Phi^* \mathcal{J} \Phi = 0$ *ist die Selbstadjungiertheit von* $a^* b$*, d. h.* $a^* b = b^* a$*.*

Definition 3.4 *Die (hermitesche) Grassmann-Mannigfaltigkeit der Lagrange-Ebenen über* $\mathfrak{SP}(2L, \mathbb{C})$ *ist gegeben durch*

$$\mathbb{L}_L^{\mathbb{C}} = \{[\Phi]_\sim : \Phi \in \mathrm{Mat}(2L \times L,\, \mathbb{C}),\, \mathrm{Rang}(\Phi) = L,\; \Phi^* \mathcal{J} \Phi = 0\}\,.$$

Beispiel 3.1 *Das einfachste Beispiel für eine Lagrange-Ebene, das auch in [Ar67] erwähnt wird, ist L* $= 1$*. In diesem Falle besteht* $\mathbb{L}_L^{\mathbb{C}} = \mathbb{L}_1^{\mathbb{C}}$ *nur aus Geraden durch den Ursprung. Die hermitesch symplektische Struktur* ω *ist dann durch die Funktion* $-\det$ *gegeben.*

Wenn $L > 1$ *ist, bildet* $\mathbb{L}_L^{\mathbb{C}}$ *eine echte Teilmenge der* L*-dimensionalen Ebenen. Betrachtet man die Ebenen* $[\Phi]_\sim$ *und* $[\widetilde{\Phi}]_\sim$*, die gegeben sind durch die Repräsentanten*

$$\Phi = \begin{pmatrix} \mathbb{1}_L \\ 0 \end{pmatrix} \qquad \text{und} \qquad \widetilde{\Phi} = \begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix},$$

so sind diese Ebenen Lagrange-Ebenen. $[\Phi]_\sim$ *wird oft als* horizontale*,* $[\widetilde{\Phi}]_\sim$ *als* vertikale *Ebene bezeichnet.*

Lemma 3.4 *Eine Lagrange-Ebene* $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ *wird durch eine Matrix* $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ *wieder auf eine Lagrange-Ebene abgebildet, d. h. ist* Φ *der Repräsentant von* $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ *und* $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$*, so gilt:*

$$\begin{aligned} \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \times \mathbb{L}_L^{\mathbb{C}} &\rightarrow \mathbb{L}_L^{\mathbb{C}} \\ (\mathfrak{M},\ \Phi) &\mapsto \mathfrak{M}\,\Phi. \end{aligned}$$

Beweis: Nach Lemma 3.3 genügt es zu zeigen, dass $(\mathfrak{M}\Phi)^* \mathcal{J} \mathfrak{M}\Phi = 0$ ist. Dazu nutzt man aus, dass $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ ist und wendet die in Lemma 3.1 vorgestellten Eigenschaften an. Es gilt:

$$(\mathfrak{M}\Phi)^* \mathcal{J} \mathfrak{M}\Phi = \Phi^* \underbrace{\mathfrak{M}^* \mathcal{J} \mathfrak{M}}_{=\mathcal{J}} \Phi = \Phi^* \mathcal{J} \Phi = 0.$$

Daraus folgt die Behauptung. □

3.3 Hermitesch symplektische und unitäre Gruppe

In diesem Abschnitt geht es darum, die Beziehungen, die zwischen der hermitesch symplektischen und der unitären Gruppe bestehen, herauszuarbeiten. Es wird sich zeigen, dass die unitäre Gruppe $\mathfrak{U}(L)$ in natürlicher Weise auf $\mathbb{L}_{\mathbb{C}}^L$ operiert und dabei die hermitesch symplektische Struktur erhält. Tatsächlich wird es sogar möglich sein, einen Diffeomorphismus zwischen den Lagrange-Ebenen und den unitären Matrizen anzugeben. Diese Tatsache wird dann ausgenutzt, um die Eigenwerte des Sturm-Liouville-Problems zu bestimmen. Der Diffeomorphismus wird aber erst im vierten Kapitel etabliert – hier geht es zunächst darum, ein intuitives Verständnis zu entwickeln.

Wir starten mit einem Beispiel, das nahelegt, dass es einen engen Zusammenhang zwischen der unitären und der hermitesch symplektischen Gruppe gibt:

Beispiel 3.2 (Unitäre Gruppe)
Seien A *und* B *in* $\mathrm{Mat}(L \times L, \mathbb{R})$*, außerdem sei* $A + iB$ *in* $\mathfrak{U}(L)$*. Da* $A + iB$ *unitär ist, folgt:*

$$\mathbb{1} = (A + iB)^*(A + iB) = A^*A + B^*B + i(A^*B - B^*A).$$

Daher gilt:

$$\begin{aligned} & A + iB \in \mathfrak{U}(L) \\ \Leftrightarrow \quad & A^*A + B^*B = \mathbb{1} \qquad \text{und} \qquad AB^* = BA^*. \end{aligned}$$

Zusammen mit (3.3) *folgt daraus, dass* $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ *in* $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ *ist.*
Die Abbildung

$$\begin{aligned} \mathrm{Mat}(L \times L,\, \mathbb{C}) &\to \mathrm{Mat}(2L \times 2L,\, \mathbb{R}) \\ A + iB &\mapsto \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \end{aligned}$$

identifiziert die unitäre Gruppe $\mathfrak{U}(L)$ *mit einer Untergruppe von* $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$.

Dieses Beispiel lässt sich verallgemeinern:

Satz 3.1 *Der Monomorphismus*

$$\mu : \mathfrak{U}(L) \to \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}),$$

der durch

$$A + iB =: \mathfrak{M} \mapsto \mu(\mathfrak{M}) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

mit A *und* $B \in \mathrm{Mat}(L \times L\, , \mathbb{R})$ *gegeben ist, identifiziert die unitäre Gruppe* $\mathfrak{U}(L)$ *mit der Untergruppe* $U(L) := \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L)$ *von* $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$.

Beweis: Ein Beweis, der dem anschließenden ähnelt, findet sich in [deG06, Kap. 2, S. 33/34], ein anderer Ansatz wurde von Arnol'd in [Ar85] gewählt.[1]

Da $\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ ist, folgt aus (3.4), dass die Inverse der Matrix $\mu(\mathfrak{M})$ durch $\begin{pmatrix} A^* & B^* \\ -B^* & A^* \end{pmatrix}$ gegeben ist. Weil A und $B \in \mathrm{Mat}(L \times L,\, \mathbb{R})$ sind, gilt:

[1] Hier sei noch ein dritter Beweis skizziert, der aber sich bereits auf die Cayley-Transformation stützt, die erst im vierten Kapitel vorgestellt wird. Man betrachte die Menge

$$\mathcal{U} = \left\{ U \in \mathfrak{U}(L) :\; U = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \right\}.$$

Nach Definition der Cayley-Transformation $\mathcal{C}$ sowie der hermitesch symplektischen, der orthogonalen und der unitären Gruppe gilt:

$$\begin{aligned} & \mathcal{C}\ \left(\mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \,\cap\, \mathfrak{O}(2L)\right) \mathcal{C}^* \\ = \ & \frac{1}{2} \begin{pmatrix} \mathbb{1}_L & -i\mathbb{1}_L \\ \mathbb{1}_L & i\mathbb{1}_L \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} \mathbb{1}_L & \mathbb{1}_L \\ i\mathbb{1}_L & -i\mathbb{1}_L \end{pmatrix} \\ = \ & \begin{pmatrix} a + ib & 0 \\ 0 & a - ib \end{pmatrix} \in \mathcal{U} \end{aligned}$$

(denn insbesondere ist $a^*b = b^*a$ und $a^*a = b^*b = \mathbb{1}$). Somit ist $\mathfrak{S}\mathfrak{P}(2L, \mathbb{R}) \,\cap\, \mathfrak{O}(2L) \cong \mathfrak{U}(L)$.

$A^* = A^T$ und $B^* = B^T$. Also ist $\big(\mu(\mathfrak{M})\big)^{-1} = \begin{pmatrix} A^T & B^T \\ -B^T & A^T \end{pmatrix} = \big(\mu(\mathfrak{M})\big)^T$. Damit ist die Gültigkeit der Inklusion $U(L) \subset \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L)$ bewiesen.

Um die andere Inklusion zu demonstrieren, betrachte man $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L)$. Da $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ ist, gilt: $\mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J}$. Darüber hinaus ist $\mathfrak{M} \in \mathfrak{O}(2L)$, d. h. $\mathfrak{M}^* = \mathfrak{M}^T$. Außerdem ist $\mathfrak{M}$ invertierbar, daraus folgt: $\mathcal{J}\mathfrak{M} = (\mathfrak{M}^T)^{-1}\mathcal{J}$. Da $\mathfrak{M}$ aber eine orthgonale Matrix ist, gilt $\mathfrak{M}^T = \mathfrak{M}^{-1}$. Insgesamt ist $\mathcal{J}\mathfrak{M} = \mathfrak{M}\mathcal{J}$, also ist $\mathfrak{M}$ in $U(L)$, bzw. $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L) \subset U(L)$. Daraus folgt die Behauptung. □

Bemerkung 3.1 *Da $\mathfrak{U}(L)$ zusammenhängend ist, folgt aus Satz 3.1 insbesondere, dass $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L)$ zusammenhängend ist.*

3.4 Lie-Gruppe und Lie-Algebra

Zunächst ist festzuhalten, dass die hermitesch symplektische Gruppe $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ eine Lie-Gruppe ist. Die zugehörige Lie-Algebra bezeichnen wir mit $\mathfrak{sp}(2L, \mathbb{C})$. Zwischen der Lie-Gruppe und der Lie-Algebra besteht eine enge Verbindung, die wir in diesem Abschnitt ausnutzen, um $\mathfrak{sp}(2L, \mathbb{C})$ näher zu charakterisieren. Darüber hinaus ist $\mathfrak{sp}(2L, \mathbb{C})$ für das Verständnis des Sturm-Liouville-Problems wichtig, da – wie wir im nächsten Abschnitt sehen werden – $\mathfrak{t}^E$ in $\mathfrak{sp}(2L, \mathbb{C})$ ist.

Satz 3.2 $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ *ist eine Lie-Gruppe.*

Beweis (vgl. [deG97, Kap. 1]): Man betrachte die stetige Abbildung

$$\begin{aligned} g:\ Gl(2L, \mathbb{C}) &\rightarrow \mathrm{Mat}(2L \times 2L,\ \mathbb{C}) \\ g(\mathfrak{M}) &= \mathfrak{M}^* \mathcal{J} \mathfrak{M} - \mathcal{J}. \end{aligned}$$

Da eine Matrix $\mathfrak{M}$ genau dann in $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ ist, wenn $\mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J}$ gilt, ist $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) = g^{-1}(0)$. Daher ist $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ eine abgeschlossene Untergruppe von $Gl(2L, \mathbb{C})$, also eine Lie-Gruppe. □

Entsprechend bildet die Gruppe aller Matrizen $\mathfrak{m} \in \mathrm{Mat}(2L \times 2L, \mathbb{C})$, für die $e^{x\mathfrak{m}} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ ist, die Lie-Algebra zu $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$. Dies ist die *hermitesch symplektische Lie-Algebra,* die – wie schon in der Einleitung zu diesem Abschnitt angekündigt – mit $\mathfrak{sp}(2L, \mathbb{C})$ bezeichnet wird. Zusammengefasst ergibt sich:

$$\mathfrak{m} \in \mathfrak{sp}(2L, \mathbb{C}) \;\Leftrightarrow\; e^{x\mathfrak{m}} =: \mathfrak{M}_x \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \text{ für alle } x \in \mathbb{R}. \tag{3.7}$$

Lemma 3.5 *Die einparametrige Familie* $(\mathfrak{M}_x)$ *aus* (3.7) *bildet eine Gruppe.*

Denn es gilt:

$$\mathfrak{M}_x \mathfrak{M}_{\tilde{x}} = e^{x\mathfrak{m}} e^{\tilde{x}\mathfrak{m}} = e^{(x+\tilde{x})\mathfrak{m}} = \mathfrak{M}_{x+\tilde{x}}$$

sowie

$$(\mathfrak{M}_x)^{-1} = (e^{x\mathfrak{m}})^{-1} = (e^{-x\mathfrak{m}}) = \mathfrak{M}_{-x}. \qquad \square$$

Satz 3.3 *Die Lie-Algebra* $\mathfrak{sp}(2L, \mathbb{C})$ *von* $\mathfrak{SP}(2L, \mathbb{C})$ *besteht aus Matrizen* $\mathfrak{m} \in \mathrm{Mat}(2L \times 2L,\ \mathbb{C})$*, die die folgenden Bedingungen erfüllen:*

$$\begin{aligned} & \mathfrak{m}\mathcal{J} + \mathcal{J}\mathfrak{m}^* = 0 \\ \text{sowie} \qquad & \mathfrak{m}^*\mathcal{J} + \mathcal{J}\mathfrak{m} = 0. \end{aligned} \tag{3.8}$$

In Blockmatrixform ausgedrückt bedeutet das:

$$\mathfrak{m} = \begin{pmatrix} A & B \\ C & -A^* \end{pmatrix} \text{ mit } A,\ B,\ C,\ D \in \mathrm{Mat}(L \times L, \mathbb{C}) \text{ und } B = B^*,\ C = C^*. \tag{3.9}$$

Beweis: Ich gehe zunächst ähnlich vor, wie in [deG06, Kap. 2, S. 36/37]. Die Gestalt in Blockmatrixform wurde von mir selbst nachgerechnet.

Zunächst zeigt man: *Alle* $\mathfrak{m} \in \mathfrak{sp}(2L, \mathbb{C})$ *erfüllen die Gleichungen* (3.8). Sei dazu $(\mathfrak{M}_x)$ eine differenzierbare, einparametrige Untergruppe von $\mathfrak{SP}(2L, \mathbb{C})$ und sei $\mathfrak{m} \in \mathrm{Mat}(2L \times 2L,\ \mathbb{C})$, so dass die Bedingung $\mathfrak{M}_x = e^{x\mathfrak{m}}$ erfüllt wird. Da $\mathfrak{M}_x$ nach Voraussetzung eine hermitesch symplektische Matrix ist, gilt: $\mathfrak{M}_x \mathcal{J} (\mathfrak{M}_x)^* = \mathcal{J}$. Setzt man für $\mathfrak{M}_x$ $e^{x\mathfrak{m}}$ ein, erhält man:

$$e^{x\mathfrak{m}} \mathcal{J} e^{x\mathfrak{m}^*} = \mathcal{J}.$$

Wenn man beide Seiten dieser Gleichung nach x ableitet, ergibt sich: $\mathfrak{m} e^{x\mathfrak{m}} \mathcal{J} e^{x\mathfrak{m}^*} + e^{x\mathfrak{m}} \mathcal{J} \mathfrak{m}^* e^{x\mathfrak{m}^*} = 0$. Setzt man $x = 0$, erhält man $\mathfrak{m}\mathcal{J} + \mathcal{J}\mathfrak{m}^* = 0$.

Da die hermitesch symplektische Gruppe nach Lemma 3.2 stabil unter der Bildung der Adjungierten ist, folgt aus $\mathfrak{M}_x \in \mathfrak{SP}(2L, \mathbb{C})$ auch, dass $(\mathfrak{M}_x)^* \in \mathfrak{SP}(2L, \mathbb{C})$ ist. Wendet man nun das gleiche Argument wie oben auf $(\mathfrak{M}_x)^*$ an, sieht man, dass ebenso $\mathfrak{m}^*\mathcal{J} + \mathcal{J}\mathfrak{m} = 0$ gilt.

Um die *Rückrichtung* zu demonstrieren, sei $\mathfrak{m}^*\mathcal{J} + \mathcal{J}\mathfrak{m} = 0$. Das Ziel ist nun zu beweisen, dass $\mathfrak{m} \in \mathfrak{sp}(2L, \mathbb{C})$ ist. Dazu genügt es zu zeigen, dass $\mathfrak{M}_x = e^{x\mathfrak{m}}$ für alle $x \in \mathbb{R}$ in $\mathfrak{SP}(2L, \mathbb{C})$ ist.

Wegen $\mathcal{J}^{-1} = -\mathcal{J}$ ist die Forderung $\mathfrak{m}^*\mathcal{J} + \mathcal{J}\mathfrak{m} = 0$ äquivalent zu $\mathfrak{m}^* = \mathcal{J}\mathfrak{m}\mathcal{J}$. Daher ist $\mathfrak{M}_x^* = e^{x\mathcal{J}\mathfrak{m}\mathcal{J}}$. Um die Reihenentwicklung der Exponentialfunktion zu bestimmen, nutzt man die bereits angesprochene Tatsache aus, dass $\mathcal{J}^2 = -\mathbb{1}$ ist und wendet dies auf $(\mathcal{J}\mathfrak{m}\mathcal{J})^k$ an. Damit ergibt sich für $k \geq 0$: $(\mathcal{J}\mathfrak{m}\mathcal{J})^k = (-1)^{k+1}\mathcal{J}\mathfrak{m}^k\mathcal{J}$. Insgesamt erhält man die Gleichung

$$e^{x\mathcal{J}\mathfrak{m}\mathcal{J}} = -\mathcal{J}\left(\sum_{k=0}^{\infty}\frac{(-x)^k}{k!}\right)\mathcal{J} = -\mathcal{J}e^{-x\mathfrak{m}}\,\mathcal{J}.$$

Daraus folgt direkt:

$$\mathfrak{M}_x^*\mathcal{J}\mathfrak{M}_x = -\mathcal{J}e^{-x\mathfrak{m}}\underbrace{\mathcal{J}\,\mathcal{J}}_{=-\mathbb{1}}e^{x\mathfrak{m}} = \mathcal{J},$$

d. h. $\mathfrak{M}_x \in \mathfrak{S}\mathfrak{P}(2L,\mathbb{C})$ wie behauptet.

Die Formel für die Schreibweise in Blockmatrizen ergibt sich, indem man zunächst $\mathfrak{m} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ setzt. Aus der Forderung $\mathfrak{m}\mathcal{J} + \mathcal{J}\mathfrak{m} = 0$ folgt:

$$\begin{aligned} 0 &= \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} + \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix} \\ &= \begin{pmatrix} B & -A \\ D & -C \end{pmatrix} + \begin{pmatrix} -B^* & -D^* \\ A^* & C^* \end{pmatrix} \\ &= \begin{pmatrix} B - B^* & -(A + D^*) \\ D + A^* & -C + C^* \end{pmatrix}. \end{aligned}$$

Somit gilt:

$$B = B^*, \quad C = C^* \quad \text{und} \quad D = -A^*.$$

Damit ist der Beweis abgeschlossen. □

3.5 Anwendung auf das Sturm-Liouville-System

Es gilt nach wie vor, das Sturm-Liouville-Problem $\mathcal{L}y = Ey$, das im zweiten Kapitel vorgestellt wurde, zu lösen. In diesem Abschnitt stelle ich den Zusammenhang zwischen dem Sturm-Liouville-Problem und der hermitesch symplektischen Geometrie her. Es wird sich zeigen, dass alle Lösungen zu einem Punkt x, die der geforderten Anfangsbedingung $u_0 \in \Psi_{\xi_0}$ genügen, Lagrange-Ebenen bilden. Um dies zu zeigen, charakterisiert man zunächst die Koeffizientenmatrix $\mathfrak{t}_x^E$.

3.5.1 Charakteristik der Koeffizientenmatrix des Sturm-Liouville-Problems

Satz 3.4 *Die Koeffizientenmatrix $\mathfrak{t}_x^E$ des Sturm-Liouville-Problems als System erster Ordnung $u'_x = \mathfrak{t}_x^E u_x$ gehört zur hermitesch symplektischen Lie-Algebra $\mathfrak{sp}(2L, \mathbb{C})$.*

Beweis: Nach der Formel (3.9) ist $\mathfrak{m} \in \mathrm{Mat}(2L \times 2L, \mathbb{C})$ genau dann in $\mathfrak{sp}(2L, \mathbb{C})$, wenn $\mathfrak{m} = \begin{pmatrix} A & B \\ C & -A^* \end{pmatrix}$ ist und wenn $B = B^*$, $C = C^*$ gilt. Für den Beweis ist es wichtig, die Voraussetzungen $v = v^*$ und $p > 0$ zu beachten. Außerdem ist wichtig, dass $E \in \mathbb{R}$ vorausgesetzt wird.

Nach (2.12) bzw. (2.13) ist

$$\mathfrak{t}^E = \begin{pmatrix} -p^{-1}q & p^{-1} \\ v - q^*p^{-1}q - E\mathbb{1}_L & q^*p^{-1} \end{pmatrix}$$

und es gilt:

$$\begin{aligned} -(q^*p^{-1})^* &= -(p^{-1})^*q \overset{p^{-1}=(p^{-1})^*}{=} -p^{-1}q \\ \text{sowie} \quad \big(v - q^*p^{-1}q - E\big)^* &= v^* - q^*(p^{-1})^*q - E = v - q^*p^{-1}q - E. \end{aligned}$$

Damit erfüllt $\mathfrak{t}^E$ alle geforderten Eigenschaften. □

Den Abschluss des Kapitels bildet der – schon mehrfach angekündigte – Satz über den Zusammenhang zwischen den Lösungen des Sturm-Liouville-Problems und den Lagrange-Ebenen.

3.5.2 Charakteristik der Lösungen des Sturm-Liouville-Problems

Satz 3.5 *Man betrachte einen Unterraum, der von den Lösungen des Systems $u'_x = \mathfrak{t}_x^E u_x$ aufgespannt wird. In Lemma 2.2 wurde gezeigt, dass jede Lösung dieses Systems auch eine Lösung von $\mathcal{L}y = Ey$ ergibt. Der Unterraum, der genau jene Lösungen umfasst, die die Anfangsbedingung $u_0 = [\Phi_0]_\sim$ erfüllen, wobei $[\Phi_0]_\sim$ eine Lagrange-Ebenen ist, bildet für jedes $x \in [0, 1]$ eine Lagrange-Ebene. Damit kann die Entwicklung der Lösungen als Entwicklung von Lagrange-Ebenen $\Phi_x \in \mathbb{L}_L^{\mathbb{C}}$ aufgefasst werden.*

Beweis: Sei u_x eine nichttriviale Lösung von $u'_x = \mathfrak{t}_x^E u_x$, d. h. $u \neq 0$, und u_x erfüllt $u'_x = \mathfrak{t}_x^E u_x$. Nach Lemma 3.3 genügt es zu zeigen, dass für alle $x \in [0, 1]$ die Bedingung $u_x^* \mathcal{J} u_x = 0$ erfüllt ist.

Es gilt:

$$\begin{aligned}
\big(u^* \mathcal{J} u\big)' &= (u')^* \mathcal{J} u + u^* \mathcal{J} u' \\
&= \big(\mathfrak{t}^E u\big)^* \mathcal{J} u + u^* \mathcal{J} \mathfrak{t}^E u \\
&= u^* \mathfrak{t}^{E*} \mathcal{J} u + u^* \mathcal{J} \mathfrak{t}^E u \\
&= u^* \underbrace{\big(\mathfrak{t}^{E*} \mathcal{J} + \mathcal{J} \mathfrak{t}^E\big)}_{=0, \text{ da } \mathfrak{t}^E \in \mathfrak{sp}(2L, \mathbb{C})} u \\
&= 0
\end{aligned} \tag{3.10}$$

Das bedeutet:

$$u_x^* \mathcal{J} u_x \equiv \text{const} \qquad \text{für alle } x \in I. \tag{3.11}$$

Nun wähle man $x = 0$. Wegen der Anfangsbedingung gilt: $u_0 = \begin{pmatrix} a \\ b \end{pmatrix} = \Phi_0^E$, wobei Φ_0^E ein Repräsentant der Ebene $[\Phi_0^E]_\sim$ sei. Somit gilt:

$$u_0^* \mathcal{J} u_0 = (\Phi_0^E)^* \mathcal{J} \Phi_0^E = 0.$$

Aufgrund von 3.11 folgt daraus für alle $x \in I$:

$$u_x^* \mathcal{J} u_x = 0,$$

und damit ist die Behauptung bewiesen. □

Bemerkung 3.2 (Anfangsbedingung)
Gemäß Definition wird das System $u'_x = \mathfrak{t}_x^E u_x$ stets unter der Anfangsbedingung $u_0 \in \Psi_{\xi_0}$ betrachtet. Darüber hinaus wurde gefordert, dass $\Psi_{\xi_0}^* \mathcal{J} \Psi_{\xi_0} = 0$ sei, d. h. $[\Psi_{\xi_0}]_\sim \in \mathbb{L}_L^{\mathbb{C}}$. Wegen Satz 3.5 sind die Lösungen des hier betrachteten Problems also tatsächlich Lagrange-Ebenen.

Bemerkung 3.3 *Es sei Φ_x der Unterraum, der alle für alle $x \in I$ die Lösungen umfasst, die die Anfangsbedingung $u_0 \in [\Phi_0]_\sim$ erfüllen. Sei weiterhin $\mathfrak{T}_x^E$ eine Fundamentalmatrix, die zu dem Problem $u'_x = \mathfrak{t}_x^E u_x$ gehört. D. h. die Spalten von $\mathfrak{T}_x^E$ bestehen aus den linear unabhängigen Lösungen zu dem hier vorgegebenen Sturm-Liouville-Problem für jeden Punkt $x \in I$. Da der Lösungsraum insgesamt*

$2L$-dimensional ist und die Fundamentalmatrix nach Definition den kompletten Lösungsraum aufspannt, gilt für die Vektoren, die die Spalten der Matrix $\mathfrak{T}_x^E$ bilden, dass sie orthogonal aufeinander stehen und eine Basis des $\mathbb{C}^{2L\times 2L}$ bilden (für alle $x \in I$). Daraus folgt sofort, dass der Rang der Matrix $\mathfrak{T}_x^E$ für alle x gleich $2L$ ist. Da $[\Phi_0]_\sim$, $[\Phi_1]_\sim$ Lagrange-Ebenen sind, folgt:

$$\text{Rang }\left([\Phi_0]_\sim\right) = L,$$

und man kann $[\Phi_0]_\sim$ durch $[\Phi_0]_\sim^\perp$ orthogonal zu einer Basis des $\mathbb{C}^{2L\times 2L}$ ergänzen.

Lemma 3.6 *Man betrachte die Abbildung $I \longrightarrow \mathfrak{sp}(2L,\mathbb{C})$ mit $x \mapsto \mathfrak{t}_x^E$. Außerdem sei $\mathfrak{T}_x^E \in \mathrm{Mat}(2L\times 2L,\ \mathbb{C})$ eine Lösung der Differentialgleichung*

$$\left(\mathfrak{T}_x^E\right)' = \mathfrak{t}_x^E \mathfrak{T}_x^E$$

mit Anfangsbedingung $\mathfrak{T}_0^E \in \mathfrak{SP}(2L,\mathbb{C})$. Dann ist $\mathfrak{T}_x^E \in \mathfrak{SP}(2L,\mathbb{C})$ für alle $x \in I$.

Beweis: Die Bedingung $\mathfrak{t}_x^E \in \mathfrak{sp}(2L,\mathbb{C})$ ist äquivalent zu $(\mathfrak{t}_x^E)^* \mathcal{J} + \mathcal{J}\,\mathfrak{t}_x^E = 0$. Also gilt:

$$\left((\mathfrak{T}_x^E)^* \mathcal{J}\,\mathfrak{T}_x^E\right)' = (\mathfrak{T}_x^E)^* \left(\mathfrak{t}_x^E\right)^* \mathcal{J}\,\mathfrak{T}_x^E + (\mathfrak{T}_x^E)^* \mathcal{J}\,\mathfrak{t}_x^E\mathfrak{T}_x^E = 0.$$

Daraus folgt:

$$(\mathfrak{T}_x^E)^* \mathcal{J}\,\mathfrak{T}_x^E = \mathfrak{T}_0^* \mathcal{J}\,\mathfrak{T}_0 = \mathcal{J},$$

und $\mathfrak{T}_x^E \in \mathfrak{SP}(2L,\mathbb{C})$ wie behauptet. □

Wie kann man Lemma 3.6 interpretieren?

Bemerkung 3.4 *Man wähle $\mathfrak{T}_0^E = \mathbb{1}_2 L$ als Anfangsbedingung der Differentialgleichung $u' = \mathfrak{t}^E u$. Dann ist nach Lemma 3.6 $\mathfrak{T}_x^E \in \mathfrak{SP}(2L,\mathbb{C})$ für alle $x \in I$. Berücksichtigt man darüber hinaus die Anfangsbedingung, die hier untersucht werden soll (nämlich $u_0 \in [\Phi_0^E]_\sim$) ergibt sich, dass die Lösungen für $x = 0$ durch $\mathfrak{T}_0^E\,[\Phi_0^E]_\sim$ bestimmt sind. Allgemeiner gilt: Da die Lösungen stets Lagrange-Ebenen sind, kann man die Lösungspfade schreiben als $[\mathfrak{T}_x^E\ \Phi_0^E]_\sim$. (Man beachte dabei, dass Lagrange-Ebenen unter hermitesch symplektischer Transformation stets auf Lagrange-Ebenen abgebildet werden.) Somit erhält man folgende Formel, um die Pfade der Lösungen zu beschreiben:*

$$[\Phi_x^E]_\sim = [\mathfrak{T}_x^E\ \Phi_0^E]_\sim, \qquad (x \in I).$$

Diese Lösungspfade sind nach wie vor analytisch in E, da die Lösungen nach Satz 5.5 für jeden Punkt x analytisch in E sind.

In diesem Kapitel stand die Gestalt der Lösungen des Sturm-Liouville-Problems im Fokus. Dazu wurde zunächst ein geeigneter Rahmen abgesteckt, der es ermöglichte, die entstehenden Strukturen zu beschreiben. Im letzten Teil wurde dann gezeigt, dass die untersuchten Lösungen Lagrange-Ebenen bilden. Wesentliche Bedingungen für dieses Verhalten waren zum einen die Randbedingungen und zum anderen die Eigenschaften der Transformationsmatrix $\mathfrak{t}_x^E$. In den nächsten Kapiteln gilt es, die hier vorgestellten Resultate zur Analyse der Eigenwerte auszunutzen.

Kapitel 4

Stereographische Projektion und Lorentz-Gruppe

Obwohl im letzten Kapitel schon einiges an Vorarbeit geleistet wurde, haben wir noch nicht genügend Handwerkszeug zusammengetragen, um die Eigenwerte des Sturm-Liouville-Problems möglichst effizient zu bestimmen. Dieses Kapitel wird uns aber einen großen Schritt weiter bringen, denn hier wird der bereits angekündigte Diffeomorphismus zwischen $\mathbb{L}_L^{\mathbb{C}}$ und $\mathfrak{U}(L)$ eingeführt und beschrieben. Dieser Diffeomorphismus wird es uns erlauben, das Sturm-Liouville-Problem unter einem neuen Blickwinkel zu betrachten, denn man kann ihn dazu nutzen, um die Lagrange-Ebenen, die, wie im vorigen Kapitel gezeigt wurde, die Lösungen des Sturm-Liouville-Problems beschreiben, auf die unitären Matrizen abzubilden.

Die Darstellung dieses Kapitels orientiert sich an [SB06].

4.1 Die stereographische Projektion

Definition 4.1 *Die* stereographische Projektion π *ist auf der folgenden Teilmenge* $\mathbb{G}_L^{inv}$ *der Grassmann-Mannigfaltigkeit* $\mathbb{G}_L^{\mathbb{C}}$ *definiert:*

$$\mathbb{G}_L^{inv} = \left\{ [\Phi]_\sim \in \mathbb{G}_L^{\mathbb{C}} : (0 \ \mathbb{1})\Phi \in Gl(L, \mathbb{C}) \right\}.$$

Ist $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ *ein Repräsentant von* $[\Phi]_\sim \in \mathbb{G}_L^{inv}$ *so ist* π *gegeben durch:*

$$\begin{aligned} \pi : \mathbb{G}_L^{inv} &\to \mathrm{Mat}(L \times L,\ \mathbb{C}) \\ \pi([\Phi]_\sim) &= \begin{pmatrix} \mathbb{1} & 0 \end{pmatrix} \Phi \left(\begin{pmatrix} 0 & \mathbb{1} \end{pmatrix} \Phi \right)^{-1} = ab^{-1}. \end{aligned}$$

Lemma 4.1 *Die stereographische Projektion π ist unabhängig vom konkret gewählten Repräsentanten.*

Denn: Seien $\Psi = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$ und $\Phi = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$ zwei Repräsentanten einer Äquivalenzklasse. Dann ist nach Definition $\Phi = \Psi\, c,\ c \in Gl(L, \mathbb{C})$, daher ist

$$\begin{aligned} \pi(\Psi) &= \pi\left(\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}\right) = \pi\left(\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} c\right) \\ &= a_1 c (b_1 c)^{-1} = a_1 b_1^{-1} = \pi(\Phi). \end{aligned} \qquad \square$$

Lemma 4.2 *Falls $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}} \cap \mathbb{G}_L^{inv}$ ist, so ist $\pi([\Phi]_\sim)$ selbstadjungiert.*

Beweis: Da $[\Phi]_\sim \in \mathbb{G}_L^{inv}$ ist, ist $\pi([\Phi]_\sim)$ wohldefiniert. Nach Lemma 4.1 ist π außerdem unabhängig vom gewählten Repräsentanten. Man wähle einen Repräsentaten $\Phi = \begin{pmatrix} a \\ b \end{pmatrix} \in [\Phi]_\sim$ mit $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$. Nach Lemma 3.3 ist $b^* a = a^* b$. Darüber hinaus ist Φ in $\mathbb{G}_L^{\mathbb{C}}$, also ist b invertierbar. Damit ergibt sich: $\pi(\Phi) = ab^{-1} = (b^{-1})^*\, a^* = (ab^{-1})^* = \big(\pi(\Phi)\big)^*$. $\square$

Die Abbildung π ist aber leider nur auf $\mathbb{G}_L^{inv}$ und nicht auf ganz $\mathbb{L}_L^{\mathbb{C}}$ definiert. Nun sucht man nach einer Methode, die es erlaubt, die Wohldefiniertheit der Projektion auf $\mathbb{L}_L^{\mathbb{C}}$ sicherzustellen. Dazu bietet es sich an, mit Hilfe der Abbildung π eine weitere Abbildung Π zu definieren. Zu diesem Zweck definiert man zunächst die *Cayley-Transformation.*

Definition 4.2 *Die* Cayley-Transformation, *bezeichnet mit $\mathcal{C}$, ist die Abbildung*

$$\mathrm{Mat}(2L \times L,\ \mathbb{C}) \ni\ \Phi \mapsto \mathcal{C}\,\Phi\ \in\ \mathrm{Mat}(2L \times L,\ \mathbb{C}).$$

Dabei bezeichnet $\mathcal{C} \in \mathrm{Mat}(2L \times 2L,\ \mathbb{C})$ die folgende Matrix:

$$\mathcal{C} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_L & -i\mathbb{1}_L \\ \mathbb{1}_L & i\mathbb{1}_L \end{pmatrix}.$$

Mit Hilfe der Cayley-Transformation kann man nun die schon angesprochene Abbildung Π definieren.

Definition 4.3 *Die Abbildung* $\Pi : \mathbb{L}_L^{\mathbb{C}} \to \mathrm{Mat}(L \times L,\ \mathbb{C})$ *ist gegeben durch:*

$$\Pi([\Phi]_\sim) = \pi\big([\mathcal{C}\,\Phi]_\sim\big) \qquad \text{für} \qquad [\mathcal{C}\,\Phi]_\sim \in \mathbb{G}_L^{inv}. \tag{4.1}$$

Das folgende Theorem beweist die Hauptresultate 1 bis 3:

Theorem 1

Die in (4.1) *definierte Abbildung* Π *ist ein reeller analytischer Diffeomorphismus von* $\mathbb{L}_L^{\mathbb{C}}$ *nach* $\mathfrak{U}(L)$.

Beweis (Idee nach [SB06]): Sei $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$. Man wählt einen Repräsentanten $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ von $[\Phi]_\sim$. Dabei sind a und $b \in \mathrm{Mat}(L \times L, \mathbb{C})$, und es gilt $a^*b = b^*a$. Weiterhin ist

$$\mathcal{C}\,\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & -i\mathbb{1} \\ \mathbb{1} & i\mathbb{1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} a - ib \\ a + ib \end{pmatrix}.$$

Nun gilt:

$$\begin{aligned} L &= \mathrm{Rang}\ (\Phi) = \mathrm{Rang}\ (\Phi^*\Phi) \\ &= \mathrm{Rang}\ \left(\begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \right) = \mathrm{Rang}\ (a^*a + b^*b) \\ &= \mathrm{Rang}\ (a^*a + i\,\underbrace{(-b^*a + a^*b)}_{=0,\ \text{da } a^*b = b^*a} + b^*b) \\ &= \mathrm{Rang}\ \big((a^* - ib^*)(a + ib)\big) = \mathrm{Rang}\ \big((a + ib)^*(a + ib)\big) \\ &= \mathrm{Rang}\ (a + ib). \end{aligned}$$

Ebenso gilt:

$$\begin{aligned} L &= \mathrm{Rang}\ (a^*a + b^*b) = \mathrm{Rang}\ \big(a^*a + i(b^*a - a^*b) + b^*b\big) \\ &= \mathrm{Rang}\ \big((a^* + ib^*)(a - ib)\big) = \mathrm{Rang}\ \big((a - ib)^*(a - ib)\big) \\ &= \mathrm{Rang}\ (a - ib). \end{aligned}$$

Insgesamt ergibt sich damit:

$$\mathrm{Rang}\ (a + ib) = \mathrm{Rang}\ (a - ib) = L \tag{4.2}$$

Es folgt, dass $[\mathcal{C}\,\Phi]_\sim \in \mathbb{G}_L^{inv}$ ist und damit im Definitionsbereich von π.

Im nächsten Schritt zeigt man, dass das Bild unter Π in $\mathfrak{U}(L)$ ist. Dazu stellt man zunächst fest, dass gilt:

$$\mathcal{C}\,\mathbb{L}_L^{\mathbb{C}} = \left\{ \left[\begin{pmatrix} a \\ b \end{pmatrix} \right]_\sim :\ a,\ b\ \in\ Gl(L, \mathbb{C}),\ a^*a = b^*b \right\}.$$

Denn: Es wurde bereits gezeigt, dass $\mathcal{C}\,\Phi = \begin{pmatrix} a - ib \\ a + ib \end{pmatrix}$ ist. Daher gilt:

$$\begin{aligned} \left(\begin{pmatrix} \mathbb{1}_L \\ 0 \end{pmatrix} \mathcal{C}\,\Phi\right)^* \begin{pmatrix} \mathbb{1}_L \\ 0 \end{pmatrix} \mathcal{C}\,\Phi &= (a-ib)^*(a-ib) \\ &= (a^* - ib^*)(a-ib) = a^*a \underbrace{-ia^*b + ib^*a}_{=0,\ da\ \ a^*b = b^*a} + b^*b \\ &= a^*a + b^*b = a^*a + \underbrace{ia^*b - ib^*a}_{=0} + b^*b \\ &= (a+ib)^*(a+ib) = \left(\begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix} \mathcal{C}\,\Phi\right)^* \begin{pmatrix} 0 \\ \mathbb{1}_L \end{pmatrix} \mathcal{C}\,\Phi. \end{aligned}$$

Es sei nun $[\mathcal{C}\,\Phi]_\sim = \left[\begin{pmatrix} a \\ b \end{pmatrix}\right]_\sim \in \mathcal{C}\,\mathbb{L}_L^{\mathbb{C}}$. Dann ist

$$\big(\Pi([\Phi]_\sim)\big)^* \Pi([\Phi]_\sim) = \left(\pi \begin{pmatrix} a \\ b \end{pmatrix}\right)^* \pi \begin{pmatrix} a \\ b \end{pmatrix} = (b^*)^{-1} a^* a b^{-1} = \mathbb{1}_L.$$

Daher gilt: $\Pi([\Phi]_\sim) \in \mathfrak{U}(L)$.

Darüber hinaus ist Π offensichtlich – als Verknüpfung von Abbildungen, die auf dem jeweiligen Definitionsbereich stetig sind – stetig auf $\mathbb{L}_L^{\mathbb{C}}$.

Die Inverse von Π gegeben ist durch

$$\Pi^{-1}(U) = \left[\begin{pmatrix} \frac{1}{2}(U + \mathbb{1}) \\ \frac{i}{2}(U - \mathbb{1}) \end{pmatrix}\right]_\sim . \tag{4.3}$$

Denn: Sei $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ ein Repräsentant von $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$. Dann gilt:

$$\begin{aligned} \Pi^{-1}\big(\Pi(\Phi)\big) &= \Pi^{-1}\big(\underbrace{(a-ib)(a+ib)^{-1}}_{\in\, \mathfrak{U}(L)} \big) \\ &= \frac{1}{2} \begin{pmatrix} \big((a-ib)(a+ib)^{-1} + \mathbb{1}_L\big) \\ i\big((a-ib)(a+ib)^{-1} - \mathbb{1}_L\big) \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \left((a-ib) + (a+ib)\right)(a+ib)^{-1} \\ i\left((a-ib) - (a+ib)\right)(a+ib)^{-1} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} (2a)(a+ib)^{-1} \\ i(-2ib)(a+ib)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} a(a+ib)^{-1} \\ b(a+ib)^{-1} \end{pmatrix}. \end{aligned}$$

Nach Voraussetzung ist $(a+ib)^{-1} \in Gl(L,\mathbb{C})$. Setzt man $c = (a+ib)^{-1}$, so gilt nach Definition von $\mathbb{L}_L^{\mathbb{C}}$, dass Φ und $\Phi\, c$ zwei Repräsentanten derselben Lagrange-Ebene sind.

Sei andererseits $U \in \mathfrak{U}(L)$, dann ist:

$$\begin{aligned}
\Pi\big(\Pi^{-1}(U)\big) &= \Pi\left(\begin{pmatrix} \frac{1}{2}(U+\mathbb{1}) \\ \frac{i}{2}(U-\mathbb{1}) \end{pmatrix}\right) \\
&= \left(\frac{1}{2}(U+\mathbb{1}) - i\frac{i}{2}(U-\mathbb{1})\right)\left(\frac{1}{2}(U+\mathbb{1}) + i\frac{i}{2}(U-\mathbb{1})\right)^{-1} = U.
\end{aligned}$$

Das heißt, Π^{-1} ist links- und rechtsinvers zu Π und damit die eindeutig definierte inverse Abbildung. Da Π^{-1} reell analytisch ist (als Verknüpfung reell analytischer Abbildungen auf dem jeweiligen Definitionsbereich), ist das Theorem damit vollständig bewiesen. □

4.2 Die Lorentz-Gruppe

Um die Cayley-Transformation gewinnbringend einzusetzen, muss man wissen, wie sich hermitesch symplektische Matrizen unter dieser Transformation verhalten. Dazu benötigt man ein weiteres wichtiges Hilfsmittel, die Lorentz-Gruppe, deren Eigenschaften dieser kurze Abschnitt gewidmet ist.

Die Lorentz-Gruppen werden hier vor allem unter einem bestimmten Blickwinkel betrachtet: Sie stellen die Gruppe der hermitesch symplektischen Matrizen $\mathfrak{SP}(2L,\mathbb{C})$ unter der Cayley-Transformation $\mathcal{C}$ dar. Darüber hinaus werde ich aber auch eine etwas allgemeinere Definition geben. Doch zunächst:

Definition 4.4 *Die Lorentz-Gruppe, $\mathcal{U}(L,L,\mathbb{C})$, ist gegeben durch:*

$$\mathcal{U}(L,L,\mathbb{C}) = \mathcal{C}\,\mathfrak{SP}(2L,\mathbb{C})\,\mathcal{C}^*. \tag{4.4}$$

In der Literatur findet man aber auch alternative Definitionen.[1] Um die obige Definition besser zu verstehen, betrachte man $\mathcal{G} \in \mathrm{Mat}(2L \times 2L, \mathbb{C})$. $\mathcal{G}$ ist – wie $\mathcal{J}$ – aus vier Blockmatrizen zusammengesetzt:

$$\mathcal{G} := \begin{pmatrix} \mathbb{1}_L & 0_L \\ 0_L & -\mathbb{1}_L \end{pmatrix}.$$

[1] Die Lorentz-Gruppe $\mathcal{U}(1,3)$ hat vielfache Anwendungen in der Relativitätstheorie hat. In diesem Falle gilt $\mathcal{U}(1,3) = \{\mathcal{T} \in Gl(4,\mathbb{R}) : \mathcal{T}\eta\mathcal{T}^T = \eta\}$. Dabei bezeichnet η die Minkowski-Metrik, d. h. $\eta = diag(-1,1,1,1)$, vgl. beispielsweise [Na03, Kap. 5, S. 208].

$\mathcal{C}$ bezeichnet weiterhin die Cayley-Transformation. Somit gilt:

$$\begin{aligned}\mathcal{C}\,\mathcal{J}\,\mathcal{C}^* &= \frac{1}{\sqrt{2}}\begin{pmatrix}\mathbb{1}_L & -i\mathbb{1}_L\\ \mathbb{1}_L & i\mathbb{1}_L\end{pmatrix}\begin{pmatrix}0 & -\mathbb{1}_L\\ \mathbb{1}_L & 0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\mathbb{1}_L & \mathbb{1}_L\\ i\mathbb{1}_L & -i\mathbb{1}_L\end{pmatrix}\\ &= \frac{1}{2}\begin{pmatrix}-2i & 0\\ 0 & 2i\end{pmatrix} = -i\begin{pmatrix}\mathbb{1}_L & 0\\ 0 & -\mathbb{1}_L\end{pmatrix} = \frac{1}{i}\,\mathcal{G}.\end{aligned} \tag{4.5}$$

Definition 4.5 *Eine alternative (aber äquivalente) Definition von* $\mathcal{U}(L, L, \mathbb{C})$ *lautet:*

$$\mathcal{U}(L, L, \mathbb{C}) = \{\mathcal{T} \in \mathrm{Mat}(2L \times 2L, \mathbb{C}) : \mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G}\}. \tag{4.6}$$

Das heißt, ähnlich wie die hermitesch symplektische Gruppe kann man auch die Lorentz-Gruppe durch eine Transformationsinvarianz definieren.

Nach Gleichung (4.5) ist $i\,\mathcal{C}\,\mathcal{J}\,\mathcal{C}^* = \mathcal{G}$. Daher ist $\mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G}$ äquivalent zu $\mathcal{T}^*\ (i\,\mathcal{C}\,\mathcal{J}\,\mathcal{C}^*)\ \mathcal{T} = i\,\mathcal{C}\,\mathcal{J}\,\mathcal{C}^*$. Erinnert man sich darüber hinaus noch an die Definition einer Matrix $\mathfrak{M} \in \mathfrak{SP}(2L, \mathbb{C})$, die sich dadurch auszeichnet, dass gilt: $\mathfrak{M}^*\,\mathcal{J}\,\mathfrak{M} = \mathcal{J}$, erschließt sich Definition 4.4.

Lemma 4.3 *Auch bei der Betrachtung der Lorentz-Gruppe ergeben sich mit der Darstellung in Blockmatrixform hilfreiche Charakterisierungen:*

$$\begin{aligned}\mathcal{U}(L, L, \mathbb{C}) &= \left\{\begin{pmatrix}A & B\\ C & D\end{pmatrix} \in Mat(2L \times 2L, \mathbb{C}) : \right.\\ &\qquad \left. A^*A - C^*C = \mathbb{1},\ D^*D - B^*B = \mathbb{1},\ A^*B = C^*D\right\}\\ &= \left\{\begin{pmatrix}A & B\\ C & D\end{pmatrix} \in Mat(2L \times 2L, \mathbb{C}) : \right.\\ &\qquad \left. A^*A - B^*B = \mathbb{1},\ D^*D - C^*C = \mathbb{1},\ AC^* = BD^*\right\}.\end{aligned} \tag{4.7}$$

Die in dieser Darstellung mit A *und* B *bezeichneten Matrizen sind invertierbar und es gilt* $\|A^{-1}B\| < 1$ *und* $\|D^{-1}C\| < 1$.

Beweis (Idee nach [SB06]): Sei $\mathcal{T} = \begin{pmatrix}A & B\\ C & D\end{pmatrix} \in \mathcal{U}(L, L, \mathbb{C})$. Nach (4.6) gilt

$\mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G}$, d. h.

$$
\begin{aligned}
&\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix}\begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix} \\
\Leftrightarrow \quad &\begin{pmatrix} A^*A - C^*C & A^*B - C^*D \\ B^*A - D^*C & B^*B - D^*D \end{pmatrix} = \begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix}.
\end{aligned}
$$

Dies ist äquivalent zum ersten Teil der Behauptung.

Der zweite Teil von Gleichung (4.7) ergibt sich dadurch, dass für $\mathcal{T} \in \mathcal{U}(L, L, \mathbb{C})$ auch $\mathcal{T}^* \in \mathcal{U}(L, L, \mathbb{C})$ ist. Denn: Sei $\mathcal{T} \in \mathcal{U}(L, L, \mathbb{C})$, dann gilt: $\mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G}$, also ist $\left(\mathcal{T}^*\mathcal{G}\mathcal{T}\right)^* = \mathcal{G}^*$. Offensichtlich ist aber $\mathcal{G}^* = \mathcal{G}$ und damit ist: $\mathcal{G} = \left(\mathcal{T}^*\mathcal{G}\mathcal{T}\right)^* = \mathcal{T}\mathcal{G}\mathcal{T}^*$, d. h. $\mathcal{T}^*$ ist tatsächlich in $\mathcal{U}(L, L, \mathbb{C})$. Bedient man sich wiederum der Schreibweise in Blockmatrixform, erhält man:

$$
\begin{aligned}
&\mathcal{T}\mathcal{G}\mathcal{T}^* = \mathcal{G} \\
\Leftrightarrow \quad &\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix}\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix} = \begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix} \\
\Leftrightarrow \quad &\begin{pmatrix} AA^* - BB^* & AC^* - BD^* \\ CA^* - DB^* & CC^* - DD^* \end{pmatrix} = \begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & -\mathbb{1}_L \end{pmatrix}.
\end{aligned}
$$

Damit ist der zweite Teil von (4.7) bewiesen.

Jetzt bleibt nur noch, die Ungleichungen und die Invertierbarkeit zu zeigen. Aus (4.7) folgt $A^*A - B^*B = \mathbb{1}$. Das impliziert, dass $\|A^*A\| \geq 1$ ist. Also ist A invertierbar. Ebenfalls aus $A^*A - B^*B = \mathbb{1}$ folgt, dass $\|A^{-1}B(A^{-1}B)^*\| = \|\mathbb{1} - A^{-1}(A^{-1})^*\| < 1$, daher ist $\|A^{-1}B\| < 1$. Das gleiche Argument, angewandt auf D und $D^{-1}C$ liefert den Beweis für die Invertierbarkeit von D bzw. dafür, dass auch $\|D^{-1}C\| < 1$ gilt. Damit ist der Beweis vollständig. □

In diesem Kapitel wurde der Diffeomorphismus zwischen $\mathbb{L}_L^{\mathbb{C}}$ und $\mathfrak{U}(L)$ etabliert. Mit Theorem 1 wurden die Hauptresultate 1–3 bewiesen. Das nächste Kapitel ist einer weiteren Verknüpfung gewidmet, der *Möbius-Transformation.*

Kapitel 5

Die Möbius-Transformation

Dieses Kapitel beschäftigt sich mit einem der letzten fehlenden Bausteine zur Bestimmung des Spektrums von $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$: mit der Möbius-Transformation. Diese Transformation dient dazu, den Zusammenhang zwischen der Lorentz-Gruppe $\mathcal{U}(L, L, \mathbb{C})$ und den Rändern des Einheitskreises $\partial\, \mathbb{D}_L^{\mathbb{C}}$ herzustellen. Dieser Rand wiederum – inbesondere der maximale Rand $\partial_L\, \mathbb{D}_L^{\mathbb{C}}$ – spielt eine bedeutende Rolle bei der Bestimmung der Eigenwerte (und ihrer jeweiligen Vielfachheit). Daher beginnt dieses Kapitel mit der Definition der oberen Halbebene, des Einheitskreises und der jeweiligen Ränder:

5.1 Obere Halbebenen und Einheitskreise

In diesem Abschnitt definiere und charakterisiere ich die Matrixdarstellungen der oberen Halbebene und des Einheitskreises. Dies ist deshalb nützlich, weil die schon vorgestellte Projektion Π die obere Halbebenen auf den Einheitskreis abbildet.

Definition 5.1 *Die obere Halbebene $\mathbb{U}_L^{\mathbb{C}}$ ist definiert durch*

$$\mathbb{U}_L^{\mathbb{C}} = \{Z \in \mathrm{Mat}(L \times L,\, \mathbb{C}) :\; i\,(Z^* - Z) > 0\},$$

während der Einheitskreis $\mathbb{D}_L^{\mathbb{C}}$ durch

$$\mathbb{D}_L^{\mathbb{C}} = \{U \in \mathrm{Mat}(L \times L,\, \mathbb{C}) :\; U^* U < \mathbb{1}\}$$

gegeben ist.[1]

[1]Es sei daran erinnert, dass $Z > 0$ für eine Matrix $Z \in \mathrm{Mat}(L \times L,\, \mathbb{C})$ bedeutet, dass Z positiv definit ist.

Der folgende Satz zeigt, dass Π via *Möbius-Transformation* die oberen Halbebenen auf die Einheitskreise abbildet. Diese Abbildung ist bijektiv.

Satz 5.1 *Die Formeln*

$$U = (Z - i\mathbb{1})(Z + i\mathbb{1})^{-1} \tag{5.1}$$

sowie

$$Z = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1} \tag{5.2}$$

beschreiben einen analytischen Diffeomorphismus von $\mathbb{U}_L^{\mathbb{C}}$ *auf* $\mathbb{D}_L^{\mathbb{C}}$.

Beweis (in Anlehnung an [Si71/73, Band III, Kap. 3, S. 122/123]): Sei $v \in \ker(Z + i\mathbb{1})$. Dann ist $(Z + i\mathbb{1})v = 0$ bzw. $iv = -Zv$. Nach der Definition der oberen Halbebene ist $i\,(Z^* - Z) > 0$ und daher gilt:[2]

$$\begin{aligned} 0 &\leq \langle v|\, i(Z^* - Z)|v\rangle \\ &= i\langle v|Z^*v\rangle + i\langle v| - Zv\rangle \\ &= i\,\langle Zv|v\rangle + i\,\langle v|iv\rangle \\ &= i\,\langle -iv|v\rangle - \,\langle v|v\rangle \\ &= -2\langle v|v\rangle. \end{aligned}$$

Daher ist $\ker(Z + i\mathbb{1}) = \{0\}$, d. h. $\det\big(Z + i\mathbb{1}\big) \neq 0$ und $(Z + i\mathbb{1})$ ist invertierbar.

Analog kann man auch die Invertierbarkeit von $\mathbb{1} - U$ überprüfen. Sei dazu wiederum $v \in \ker(\mathbb{1} - U)$. Somit ist $(\mathbb{1} - U)v = 0$, d. h. $v = Uv$. Damit gilt:

$$\langle v|\mathbb{1} - U^*U|v\rangle = \langle v|v\rangle - \langle Uv|Uv\rangle = \langle v|v\rangle - \langle v|v\rangle = 0.$$

Da nach Voraussetzung $\mathbb{1} > U^*U$ bzw. $\mathbb{1} - U^*U > 0$ ist, ist $v = 0$. Daher ist $\det\big(\mathbb{1} - U\big) \neq 0$ und die Inverse $(\mathbb{1} - U)^{-1}$ existiert.

Sowohl $\mathbb{1} + U$ und $\mathbb{1} - U$ als auch $Z - i\mathbb{1}$ und $Z + i\mathbb{1}$ kommutieren, denn $(\mathbb{1} + U)(\mathbb{1} - U) = \mathbb{1} + U - U + U^2 = \mathbb{1} - U + U + U^2 = (\mathbb{1} - U)(\mathbb{1} + U)$ und genauso: $(Z - i\mathbb{1})(Z + i\mathbb{1}) = Z^2 + iZ - iZ + \mathbb{1} = (Z + i\mathbb{1})(Z - i\mathbb{1})$. Daher erhalten beide Abbildungen die Symmetrie.

[2] An dieser Stelle sei daran erinnert, dass das Skalarprodukt $\langle\,.\,|\,.\,\rangle$ linear im zweiten und antilinear im ersten Argument ist.

Zum Abschluss rechnet man nun nach, dass (5.1) und (5.2) Inverse voneinander sind. Zum einen gilt für $u \in \mathbb{U}_L^{\mathbb{C}}$:

$$\begin{aligned} U\big(Z(u)\big) &= U\Bigg(\underbrace{i\,\frac{\mathbb{1}+u}{\mathbb{1}-u}}_{\in\,\mathbb{D}}\Bigg) = \left(i\,\frac{\mathbb{1}+u}{\mathbb{1}-u} - i\mathbb{1}\right)\left(i\,\frac{\mathbb{1}+u}{\mathbb{1}-u} + i\mathbb{1}\right)^{-1} \\ &= \frac{\mathbb{1}+u-\mathbb{1}+u}{\mathbb{1}-u}\;\frac{\mathbb{1}-u}{\mathbb{1}+u+\mathbb{1}-u} = u. \end{aligned}$$

Zum anderen ergibt sich für $z \in \mathbb{D}_L^{\mathbb{C}}$:

$$\begin{aligned} Z\big(U(z)\big) &= Z\Bigg(\underbrace{\frac{z-i\mathbb{1}}{z+i\mathbb{1}}}_{\in\,\mathbb{U}}\Bigg) = i\left(\mathbb{1}+\frac{z-i\mathbb{1}}{z+i\mathbb{1}}\right)\left(\mathbb{1}-\frac{z-i\mathbb{1}}{z+i\mathbb{1}}\right)^{-1} \\ &= i\,\frac{z+i\mathbb{1}+z-i\mathbb{1}}{z+i\mathbb{1}}\;\frac{z+i\mathbb{1}}{-z+i\mathbb{1}+z+i\mathbb{1}} = i\,\frac{2z}{2i\mathbb{1}} = z. \end{aligned}$$

Damit sind alle Behauptungen bewiesen. □

Definition 5.2 (Ränder von $\mathbb{U}_L^{\mathbb{C}}$ und $\mathbb{D}_L^{\mathbb{C}}$)
Der Rand $\partial\mathbb{U}_L^{\mathbb{C}}$ der oberen Halbebene $\mathbb{U}_L^{\mathbb{C}}$ ist ein stratifizierter Raum, der aus der Vereinigung der Strata $\partial_l\mathbb{U}_L^{\mathbb{C}}$, $(l = 1, \ldots, L)$ besteht. Dabei ist $\partial_l\mathbb{U}_L^{\mathbb{C}}$ die Menge der Matrizen Z, für die $i(Z^* - Z)$ vom Rang $L - l$ ist.

So ähnlich geht man auch bei der Definition des Randes der Einheitskreisscheibe $\mathbb{D}_L^{\mathbb{C}}$ vor. Der Rand $\partial\mathbb{D}_L^{\mathbb{C}}$ ist wiederum ein stratifizierter Raum, diesmal sind die Strata gegeben durch $\partial_l\mathbb{D}_L^{\mathbb{C}}$ $(l = 1, \ldots, L)$. Jedes Stratum besteht aus den Matrizen U, die die Bedinungen $U^*U \leq \mathbb{1}$ und $\mathrm{Rang}(\mathbb{1} - U^*U) = L - l$ erfüllen.

Bemerkung 5.1 (Maximale Ränder von $\mathbb{U}_L^{\mathbb{C}}$ und $\mathbb{D}_L^{\mathbb{C}}$)
Die hermiteschen Matrizen stellen den maximalen Rand der oberen Halbebene $\partial_L\mathbb{U}_L^{\mathbb{C}}$ dar, denn für diese Matrizen gilt $Z^* = Z$, daher ist $i(Z^* - Z)$ vom Rang $0 = L - L$.

Auf ähnliche Weise erhält man für den maximalen Rand der Einheitskreise:

$$\partial_L\mathbb{D}_L^{\mathbb{C}} = \mathfrak{U}(L). \tag{5.3}$$

Denn da für die unitären Matrizen $U^* = U^{-1}$ gilt, erhält man: $\mathrm{Rang}(\mathbb{1} - U^*U) = \mathrm{Rang}(\mathbb{1} - \mathbb{1}) = 0 = L - L$. *Nach Theorem 1 kann der maximale Rand $\partial_L\mathbb{D}_L^{\mathbb{C}}$ über die Abbildung Π mit $\mathbb{L}_L^{\mathbb{C}}$ identifiziert werden.*

Darüber hinaus ist zu bemerken, dass die Abbildung Π am Rand Singularitäten hat und die verschiedenen Strata vermischt. Insbesondere wird $\partial_L \mathbb{U}_L^{\mathbb{C}}$ nicht auf $\partial_L \mathbb{D}_L^{\mathbb{C}}$ abgebildet.

5.2 Die Möbius-Transformation

Dieser Abschnitt dient dazu, die Möbius-Transformation einzuführen und ihre Wohldefiniertheit zu belegen. Ihre Eigenschaften stehen dann im nächsten Abschnitt im Mittelpunkt der Untersuchung. Zur Definition der Möbius-Transformation ist es wiederum sinnvoll, sich der Schreibweise in Blockmatrizen zu bedienen.

Definition 5.3 (Möbius-Transformation)
Seien $\mathcal{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Gl(2L, \mathbb{C})$ *und* $Z \in \mathrm{Mat}(L \times L, \mathbb{C})$. *Sei ferner* $(CZ + D)$ *invertierbar. Dann ist die* Möbius-Transformation[3] *definiert durch:*

$$\mathcal{T} \cdot Z = (AZ + B)\,(CZ + D)^{-1}. \tag{5.4}$$

Nun ist es möglich, über die stereographische Projektion π eine Verbindung zu den Lagrange-Ebenen herzustellen:

Lemma 5.1 *Seien* $[\Phi]_\sim$ *und* $[\mathcal{T}\Phi]_\sim \in \mathbb{G}_L^{inv}$, π *sei die stereographische Projektion. Dann definiert die Möbius-Transformation eine Matrixmultiplikation:*

$$\pi\big(\mathcal{T}([\Phi]_\sim)\big) = \mathcal{T} \cdot \pi([\Phi]_\sim). \tag{5.5}$$

Beweis: Zunächst zur Notation, es sei $\mathcal{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ und $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$ ein Repräsentant von $[\Phi]_\sim$. Dann ist $\pi(\Phi) = ab^{-1}$. Im ersten Beweisschritt ist sicherzustellen, dass alle Ausdrücke *wohldefiniert* sind, d. h. dass alle benötigten Inversen existieren. $[\mathcal{T}\,\Phi]_\sim \in \mathbb{G}_L^{inv}$ impliziert, dass $Ca + Db$ invertierbar ist. Da $[\Phi]_\sim \in \mathbb{G}_L^{inv}$ ist, ist b invertierbar. Daraus folgt, dass auch $Cab^{-1} + D = C\pi([\Phi]_\sim) + D$ invertierbar ist. Damit ist die Möbius-Transformation $\mathcal{T} \cdot \pi([\Phi]_\sim)$ wohldefiniert.

[3] Die Möbius-Transformation wird in der Literatur manchmal auch kanonische Transformation oder fraktale Transformation genannt.

Im zweiten Schritt berechnet man:

$$\begin{aligned}\pi\big(\mathcal{T}(\Phi)\big) &= \pi\left(\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\right) = \pi\left(\begin{pmatrix} Aa+Bb \\ Ca+Db \end{pmatrix}\right) \\ &= \big(Aa+Bb\big)\big(Ca+Db\big)^{-1} = \big(Aab^{-1}+B\big)\big(Cab^{-1}+D\big)^{-1} \\ &= \mathcal{T}\cdot\pi(\Phi)\,.\end{aligned}$$

Damit ist das Lemma bewiesen. □

Bemerkung 5.2 *Es seien* $\big(\mathcal{C}\mathcal{T}[\Phi]_\sim\big)$ *und* $\big(\mathcal{C}[\Phi]_\sim\big)$ *in* $\mathbb{G}_L^{inv}$*. Beachtet man, dass* $\mathcal{C}^* = \mathcal{C}^{-1}$ *ist, so folgt aus* (5.5)*:*

$$\pi\big(\mathcal{C}\mathcal{T}[\Phi]_\sim\big) = \pi\big(\mathcal{C}\mathcal{T}\mathcal{C}^*\ \mathcal{C}[\Phi]_\sim\big) = \mathcal{C}\mathcal{T}\mathcal{C}^*\cdot\Pi([\Phi]_\sim). \tag{5.6}$$

Im folgenden Fall sind die Bedingungen, die in (5.5) gefordert wurden, um die Wohldefiniertheit zu garantieren, automatisch erfüllt.

Satz 5.2
Die hermitesch symplektische Gruppe $\mathfrak{S}\mathfrak{P}(2L,\mathbb{C})$ *operiert mittels Möbius-Transformation auf* $\mathbb{U}_L^{\mathbb{C}}$*.*
Die Lorentz-Gruppe $\mathcal{U}(L,L,\mathbb{C})$ *operiert durch die Möbius-Transformation auf* $\mathbb{D}_L^{\mathbb{C}}$*.*

Beweis: Dieser Satz kann wie kann wie bei [Si71/73] bewiesen werden.

5.3 Eigenschaften der Möbius-Transformation

Natürlich ist nicht nur die bloße Existenz der Möbius-Transformation von Interesse, sondern man möchte ihre Eigenschaften verstehen und ausnutzen. Diesen Eigenschaften ist der folgende Abschnitt gewidmet. Im Anschluss werden wir uns den Anwendungen auf Sturm-Liouville-Probleme zuwenden.

Der folgende Satz belegt, dass die Lorentz-Gruppe durch Möbius-Transformation nicht nur auf dem Einheitskreis, sondern auch auf dessen Rand $\partial\,\mathbb{D}_L^{\mathbb{C}}$ (der in Strata $\partial_l\,\mathbb{D}_L^{\mathbb{C}}$ zerfällt) operiert. Die Möbius-Transformation auf dem maximalen Rand $\partial_L\,\mathbb{D}_L^{\mathbb{C}}$ implementiert die Operation der hermitesch symplektischen Gruppe $\mathfrak{S}\mathfrak{P}(2L,\mathbb{C})$ auf der Grassmann-Mannigfaltigkeit der Lagrange-Ebenen $\mathbb{L}_L^{\mathbb{C}}$. Wegen der schon angesprochenen Singularitäten kann man aber keine ähnliche Aussage für den Rand $\partial\,\mathbb{U}_L^{\mathbb{C}}$ der oberen Halbebene (oder eines ihrer Strata $\partial_l\mathbb{U}_L^{\mathbb{C}}$) treffen.

Satz 5.3 *Die Lorentz-Gruppe* $\mathcal{U}(L,L,\mathbb{C})$ *operiert via Möbius-Transformation auf* $\partial_l \mathbb{D}_L^{\mathbb{C}}$. *Im Falle* $l = L$ *implementiert die Möbius-Transformation eine Operation von* $\mathfrak{SP}(2L,\mathbb{C})$ *auf* $\mathbb{L}_L^{\mathbb{C}}$*: Sei* $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ *und* $\mathfrak{T} \in \mathfrak{SP}(2L,\mathbb{C})$*, dann gilt:*

$$\Pi([\mathfrak{T}\Phi]_\sim) = \mathcal{C}\ \mathfrak{T}\ \mathcal{C}^*\ \cdot\ \Pi\big([\Phi]_\sim\big).$$

Beweis (Idee nach [SB06]): Wegen (5.6) muss man nur noch zeigen, dass für $U \in \partial_l \mathbb{D}_L^{\mathbb{C}}$ und $\mathcal{T},\ \mathcal{T}_1,\ \mathcal{T}_2 \in \mathcal{U}(L,L,\mathbb{C})$ die Möbius-Transformation $\mathcal{T}\cdot U$ erstens wohldefiniert und zweitens wieder in $\partial_l\ \mathbb{D}_L^{\mathbb{C}}$ ist. Drittens ist zu zeigen, dass die Beziehung $(\mathcal{T}_1\mathcal{T}_2)\cdot U = \mathcal{T}_1 \cdot (\mathcal{T}_2 \cdot U)$ gilt.

1. Zur *Wohldefiniertheit:* Sei $\mathcal{T} \in \mathcal{U}(L,L,\mathbb{C})$: Für $A,\ B,\ C$ und $D \in \mathrm{Mat}(L\times L,\ \mathbb{C})$ sei $\mathcal{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, wobei die Eigenschaften aus Lemma 4.3 gelten.

   Dann besagt eben dieses Lemma, dass $\|D^{-1}C\| < 1$ ist. Das impliziert die Invertierbarkeit von $(CU + D) = D\,(\mathbb{1} + D^{-1}CU)$, somit ist die Möbius-Transformation wohldefiniert.

2. Nun gilt es zu zeigen, dass $\mathcal{T}\cdot U \in \partial_l\ \mathbb{D}_L^{\mathbb{C}}$ ist. Um dies zu demonstrieren, greift man wieder auf Lemma 4.3 zurück. Damit gilt:

$$\begin{aligned}
&(CU+D)^*(CU+D) - (AU+B)^*(AU+B) \\
&= U^*C^*CU + U^*C^*D + D^*CU + D^*D \\
&\ - \big(U^* \underbrace{A^*A}_{=C^*C+\mathbb{1}} U + U^* \underbrace{A^*B}_{=C^*D} + \underbrace{B^*A}_{=D^*C} U + \underbrace{B^*B}_{D^*D-\mathbb{1}}\big) \\
&= \mathbb{1} - U^*U.
\end{aligned} \tag{5.7}$$

   Jetzt mulitpliziere man (5.7) von links mit $\big((CU+D)^*\big)^{-1}$ und von rechts mit $(CU+D)^{-1}$. Damit ergibt sich:

$$\begin{aligned}
&\big((CU+D)^*\big)^{-1}\Big((CU+D)^*(CU+D) - (AU+B)^*(AU+B)\Big)(CU+D)^{-1} \\
&= \big((CU+D)^*\big)^{-1}(\mathbb{1} - U^*U)(CU+D)^{-1}.
\end{aligned}$$

   Das ist äquivalent zu

$$\begin{aligned}
&\mathbb{1} - \big((AU+B)(CU+D)^{-1}\big)^*(AU+B)(CU+D)^{-1} \\
&= ((CU+D)^*)^{-1}(\mathbb{1} - U^*U)(CU+D)^{-1}.
\end{aligned} \tag{5.8}$$

Nach der Definition 5.3 ist $\mathcal{T} \cdot U = (AU + B)\,(CU + D)^{-1}$. Setzt man dies in Formel (5.8) ein, erhält man

$$\begin{aligned} & (\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U) \\ = \;& \mathbb{1} - \big((CU + D)^*\big)^{-1}(\mathbb{1} - U^*U)(CU + D)^{-1}. \end{aligned} \tag{5.9}$$

Nun benötigt man einige Abschätzungen: Für $U \in \partial_l \mathbb{D}_L^{\mathbb{C}}$ gilt $\mathbb{1} - U^*U \geq 0$. Wie man schon beim Beweis der Wohldefiniertheit ausgenutzt hat, liefert Lemma 4.3: $\|D^{-1}C\| < 1$. Daher ist $(CU + D) = D\,(\mathbb{1} + D^{-1}CU) > 0$. Dies impliziert, dass auch $(CU + D)^{-1} > 0$ und ebenso $((CU + D)^*)^{-1} > 0$ gilt. Setzt man dies in (5.9) ein, ergibt sich die gesuchte Abschätzung:

$$(T \cdot U)^*(T \cdot U) \leq \mathbb{1}.$$

Als nächstes zeigen wir, dass die Matrix $(CU+D)$ (die nach Lemma 4.3 und obigen Bemerkungen zur Wohldefiniertheit invertierbar ist) $\ker(\mathbb{1} - U^*U)$ auf $\ker(\mathbb{1} - (\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U))$ abbildet. Außerdem wird $\ker(\mathbb{1} - U^*U)^\perp$ unter $(CU + D)$ auf $\ker(\mathbb{1} - (\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U))^\perp$ abgebildet.

Um dies zu demonstrieren, gibt man sich ein $v \in \ker(\mathbb{1} - U^*U)$ vor. Gleichung (5.7) impliziert, dass

$$\begin{aligned} \big(\mathbb{1} - U^*U\big)v &= 0 \\ \Leftrightarrow \quad \big((CU + D)^*(CU + D) - (AU + B)^*(AU + B)\big)v &= 0. \end{aligned}$$

Daraus ergibt sich:

$$\begin{aligned} \|(CU + D)\,v\| &= \|(AU + B)\,v\| \\ &= \big\|(AU + B)(CU + D)^{-1}(CU + D)\,v\big\| \\ &= \|T \cdot U\,(CU + D)\,v\|\,. \end{aligned}$$

Da $\mathbb{1} - \big((\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U)\big) \geq 0$ ist, kann man aus der obigen Gleichung folgendes schließen:

$$\begin{aligned} 0 &\leq \langle (CU + D)\,v | \mathbb{1} - \big((\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U)\big) | (CU + D)\,v \rangle \\ &= \langle (CU + D)\,v | (CU + D)\,v \rangle \\ &\quad - \langle (CU + D)\,v | \big((\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U)\big)(CU + D)\,v \rangle \\ &= \langle (CU + D)\,v | (CU + D)\,v \rangle \\ &\quad - \langle (\mathcal{T} \cdot U)\,(CU + D)\,v | (\mathcal{T} \cdot U)\,(CU + D)\,v \rangle \\ &= \|(CU + D)\,v\| \;-\; \|(\mathcal{T} \cdot U)\,(CU + D)\,v\| \\ &= \|(CU + D)\,v\| \;-\; \|(CU + D)\,v\| \\ &= 0 \end{aligned}$$

Somit ist $(CU + D)\, v \in \ker\big(\mathbb{1} - (\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U)\big)$.

Ähnlich geht man auch für $\tilde{v} \in \ker(\mathbb{1} - U^*U)^\perp$ vor: $\tilde{v} \in \ker(\mathbb{1} - U^*U)^\perp$ impliziert $\big(\mathbb{1} - U^*U\big)\tilde{v} > 0$. Zunächst folgert man aus (5.7) wie oben:

$$\big((CU + D)^*(CU + D) - (AU + B)^*(AU + B)\big)\tilde{v} > 0.$$

Somit gilt:

$$\begin{aligned} \|(CU + D)\, v\| &> \|(AU + B)\, \tilde{v}\| \\ &= \|\mathcal{T} \cdot U\, (CU + D)\, \tilde{v}\|\, . \end{aligned}$$

Daraus folgert man analog zur Rechnung für $v \in \ker(\mathbb{1} - U^*U)$, dass für $\tilde{v} \in \ker(\mathbb{1} - U^*U)^\perp$ gilt:

$$|(CU + D)\tilde{v}| > |\mathcal{T} \cdot U(CU + D)\tilde{v}|\, .$$

Daraus schließt man, dass $\tilde{v} \notin \ker(\mathbb{1} - (\mathcal{T} \cdot U)^*(\mathcal{T} \cdot U))$ ist.

3. Man rechnet nach, dass tatsächlich gilt $(\mathcal{T}_1\, \mathcal{T}_1) \cdot Z = \mathcal{T}_1 \cdot (\mathcal{T}_2 \cdot Z)$:

$$\begin{aligned} (\mathcal{T}_1\, \mathcal{T}_1) \cdot Z &= \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} \begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} \cdot Z \\ &= \begin{pmatrix} A_1A_2 + B_1C_2 & A_1B_2 + B_1D_2 \\ C_1A_2 + D_1C_2 & C_1B_2 + D_1D_2 \end{pmatrix} \cdot Z \\ &= ((A_1A_2 + B_1C_2)Z + A_1B_2 + B_1D_2) \\ &\quad ((C_1A_2 + D_1C_2)Z + C_1B_2 + D_1D_2)^{-1} \end{aligned}$$

Andererseits ist

$$\begin{aligned} \mathcal{T}_1 \cdot (\mathcal{T}_2 \cdot Z) &= \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} \cdot \left(\begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} \cdot Z \right) \\ &= \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} \cdot (A_2Z + B_2)(C_2Z + D_2)^{-1} \\ &= (A_1(A_2Z + B_2)(C_2Z + D_2)^{-1} + B_1) \\ &\quad (C_1(A_2Z + B_2)(C_2Z + D_2)^{-1} + D_1)^{-1} \\ &= (A_1(A_2Z + B_2) + B_1(C_2Z + D_2)) \\ &\quad ((C_1(A_2Z + B_2)(C_2Z + D_2)^{-1} + D_1)(C_2Z + D_2))^{-1} \\ &= ((A_1A_2 + B_1C_2)Z + A_1B_2 + B_1D_2) \\ &\quad ((C_1A_2 + D_1C_2)Z + C_1B_2 + D_1D_2)^{-1} \end{aligned}$$

Damit ist der Beweis abgeschlossen. □

5.4 Projektion des Sturm-Liouville-Problems auf den Raum der unitären Matrizen

Im letzten Abschnitt dieses Kapitels steht die Frage im Mittelpunkt, wie man die Möbius-Transformation auf die Lösungen der Sturm-Liouville-Probleme, die in Kapitel 2 vorgestellt wurden, anwenden kann. Es sei daran erinnert, dass nach Lemma 3.6 und Bemerkung 3.4 die Pfade der Lösungen des Sturm-Liouville-Problems durch Pfade von Lagrange-Ebenen gegeben sind. Diese Pfade können durch $[\Phi_x^E]_\sim = [\mathfrak{T}^E \Phi_0^E]_\sim$, $\mathfrak{T}_x^E \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ und $[\Phi_0^E]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ beschrieben werden.

Satz 5.4 (Neuformulierung des Eigenwertproblems)
Die Lösungen des Sturm-Liouville-Problems stellen sich unter der Möbius-Transformation wie folgt dar:

$$\Pi\big([\Phi_x^E]_\sim\big) = \mathcal{C}\, \mathfrak{T}_x^E\, \mathcal{C}^* \cdot \Pi([\Phi_0^E]_\sim).$$

Beweis: Da $[\Phi_0^E]_\sim$ und $\mathfrak{T}^E\, [\Phi_0^E]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ sind, gilt nach Theorem 1:

$$[\mathcal{C}\Phi_0^E]_\sim \in \mathbb{G}_L^{inv} \quad \text{und} \quad [\mathcal{C}\mathfrak{T}^E\, \Phi_0^E]_\sim = [\mathcal{C}\mathfrak{T}^E\mathcal{C}^*\mathcal{C}\Phi_0^E]_\sim \in \mathbb{G}_L^{inv}.$$

Daher ist

$$\Pi\left([\mathfrak{T}^E\, \Phi_0^E]_\sim\right) = \pi\left([\mathcal{C}\mathfrak{T}^E\, \Phi_0^E]_\sim\right) = \mathcal{C}\mathfrak{T}^E\mathcal{C}^* \cdot \pi\left([\Phi_0^E]_\sim\right). \tag{5.10}$$

Aus der Definition von Π folgt die Behauptung. □

Satz 5.5 (Analytizität in *E*)
Das System (5.10) *ist analytisch in E.*

Beweis: Um diesen Satz zu beweisen, muss man nicht mehr viel arbeiten; es genügt, die bisherigen Resultate richtig zusammenzusetzen. Die Möbius-Transformation ist – auf dem hier gewählten Definitionsbereich – eine analytische Funktion. Die Analytizität von (5.10) ergibt sich direkt aus der Analytizität des ursprünglichen Problems (2.12), das nach Satz 2.4 analytisch in E ist. Daher sind die Lösungen des Problems Φ_x^E analytisch in E. Da die Matrizen $\mathfrak{T}_x^E$ und die Lagrange-Ebene $[\Phi_0^E]_\sim$ dazu dienen, Pfade dieser Lösungen zu beschreiben, ist auch $[\mathfrak{T}_x^E\, \Phi_0^E]_\sim = [\Phi_x^E]_\sim$ analytisch in E. Damit entsteht $\Pi\big([\mathfrak{T}^E(x)\, \Phi_0^E]_\sim\big)$ durch die Verknüpfung analytischer Funktionen. Da die Verknüpfung mehrerer analytischer Funktionen aber wieder eine analytische Funktion ist, folgt die Behauptung. □

Satz 5.6 (Differentialgleichung auf $\mathfrak{U}(L)$)
Transformiert man die Lösungen des Sturm-Liouville-Problems mit Hilfe der stereographischen Projektion auf $\mathfrak{U}(L)$ so ergeben sich die resultierenden unitären Matrizen als Lösungen der (nichtlinearen) Differentialgleichung

$$\partial_x\, U_x^E = \begin{pmatrix} \mathbb{1} \\ -(U_x^E)^* \end{pmatrix}^* \mathcal{C}\, \mathfrak{t}_x^E\, \mathcal{C}^* \begin{pmatrix} U_x^E \\ \mathbb{1} \end{pmatrix}. \tag{5.11}$$

Ist die Anfangsbedingung des Problems $u_x' = \mathfrak{t}_x^E u_x$ gegeben durch $u_0 = [\Phi_0^E]_\sim$, so ist die Anfangsbedingung des transformierten Problems $\Pi\left(\Phi_0^E\right)$, wobei Φ_0^E ein Repräsentant von $[\Phi_0^E]_\sim$ ist.

Beweis: Die Lösungen des Sturm-Liouville-Problem zu den Anfangsbedingungen $u_0 = [\Phi_0^E]_\sim$ sind nach Satz 3.5 Lagrange-Ebenen. Daher kann man sich darauf beschränken, die Differentialgleichung $\partial_x[\Phi_x^E]_\sim = [\mathfrak{t}_x^E\ \Phi_x^E]_\sim$ zu untersuchen. Wie üblich schreibt man $\Phi = \begin{pmatrix} a^* & b^* \end{pmatrix}^*$. Wendet man die Cayley-Transformation darauf an, erhält man:

$$\begin{aligned} \mathcal{C}\,\partial_x[\Phi_x^E]_\sim = \partial_x[\mathcal{C}\,\Phi_x^E]_\sim &= [\mathcal{C}\,\mathfrak{t}_x^E\,\Phi_x^E]_\sim = \mathcal{C}\,\mathfrak{t}_x^E\,\mathcal{C}^*[\mathcal{C}\,\Phi_x^E]_\sim \\ \Leftrightarrow \qquad \partial_x \begin{pmatrix} a_x - ib_x \\ a_x + ib_x \end{pmatrix} &= \mathcal{C}\mathfrak{t}_x^E\mathcal{C}^* \begin{pmatrix} a_x - ib_x \\ a_x + ib_x \end{pmatrix}. \end{aligned}$$

Nach Theorem 1 ist die Abbildung $\pi\left(\mathcal{C}\Phi_x^E\right)$ wohldefiniert und in $\mathfrak{U}(L)$. Somit folgt:

$$\begin{aligned} \partial_x\, U_x^E &= \partial_x\big((a_x - ib_x)(a_x + ib_x)^{-1}\big) \\ &= \partial_x(a_x - ib_x)\ (a_x + ib_x)^{-1} + (a_x - ib_x)\ \partial_x\big((a_x + ib_x)^{-1}\big) \\ &= \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix}^* \mathcal{C}\mathfrak{t}_x^E\mathcal{C}^* \begin{pmatrix} a_x - ib_x \\ a_x + ib_x \end{pmatrix} (a_x + ib_x)^{-1} \\ &\quad -(a_x - ib_x)(a_x + ib_x)^{-1}\partial_x(a_x + ib_x)\ (a_x + ib_x)^{-1} \\ &= \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix}^* \mathcal{C}\mathfrak{t}_x^E\mathcal{C}^* \begin{pmatrix} U_x^E \\ \mathbb{1} \end{pmatrix} - U_x^E \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix}^* \mathcal{C}\mathfrak{t}_x^E\mathcal{C}^* \begin{pmatrix} U_x^E \\ \mathbb{1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{1} \\ (-U_x^E)^* \end{pmatrix}^* \mathcal{C}\mathfrak{t}_x^E\mathcal{C}^* \begin{pmatrix} U_x^E \\ \mathbb{1} \end{pmatrix}. \end{aligned}$$

Da sich die Anfangsbedingung durch schlichtes Einsetzen in die Formel ergibt, ist die Behauptung bewiesen. □

Kapitel 6

Der Bott-Maslov-Index

Der Index, der in diesem Kapitel im Mittelpunkt steht, wird uns in die Lage versetzen, die schon mehrfach angekündigten Aussagen über das Verhalten der Eigenwerte von Sturm-Liouville-Problemen zu beweisen. Bott-Maslov-Indizes findet man häufig in der Literatur. Sie bezeichnen Funktionen, die im Allgemeinen auf Pfaden in $\mathbb{L}_L^{\mathbb{C}}$ definiert sind[1] und nur diskrete Werte annehmen. Wozu dient eine solche Funktion? Der Bott-Maslov-Index wird genutzt, um die Anzahl der Schnitte eines Pfades in $\mathbb{L}_L^{\mathbb{C}}$ mit einem singulären Zykel[2], die ebenfalls in $\mathbb{L}_L^{\mathbb{C}}$, aber fest gewählt ist, zu beschreiben.

Da der Index, dem dieses Kapitel der Arbeit gewidmet ist, in der Literatur unter dem Namen Maslov-Index firmiert,[3] werde ich auch in dieser Arbeit den Namen Maslov zitieren, um den Leser nicht zu irritieren. Es sei aber darauf hingewiesen, dass Bott (in [Bo56]) diesen Index schon früher als Maslov beschrieb – um Gerechtigkeit walten zu lassen, nenne ich das als „Maslov-Index" bekannte Objekt „Bott-Maslov-Index".

Die hier gewählte Definition des Bott-Maslov-Indexes sowie die Gliederung des Kapitels orientiert sich an [SB06].

[1]Man findet aber auch Autoren, die den Index für Pfade oder Zykel in $\mathfrak{SP}(2L, \mathbb{C})$ definieren (vgl. [deG06, Kap. 3.1, S. 67–74]).

[2]Diese kritische Stelle bezeichnet man auch als *Kaustik,* vgl. [CLR90].

[3]Da auch Arnol'd einen großen Teil zum Verständnis dieses Indexes beitrug, findet man in der Literatur teilweise die Bezeichung „Arnol'd-Maslov-Index".

6.1 Singuläre Zykel

Als erstes beschäftigt man sich mit den kritischen Stellen, die später die Schnittstellen bilden werden. Dazu definiert man einen Zykel, der den nicht trivialen Schnitt von Lagrange-Ebenen mit einer gegebenen Ebene $[\Psi]_\sim$ beschreibt.

Definition 6.1 *Es sei* $\Psi_\xi = \begin{pmatrix} \xi \\ \mathbb{1} \end{pmatrix}$ *mit* $\xi \in \mathrm{Her}(L, \mathbb{C})$*. Der zu* $[\Psi_\xi]_\sim$ *assoziierte singuläre Zykel ist gegeben durch:*

$$\mathbb{L}_L^{\mathbb{C},\,\xi} = \Big\{ [\Phi]_\sim \,\in\, \mathbb{L}_L^{\mathbb{C}} :\; \Phi\, \mathbb{C}^L \,\cap\, \Psi_\xi\, \mathbb{C}^L \,\neq\, \{\overrightarrow{0}\} \Big\}. \tag{6.1}$$

Diesen singulären Zykel bezeichnet man als Bott-Maslov-Zykel.[4] *Er kann in eine disjunkte Vereinigung* $\mathbb{L}_L^{\mathbb{C},\,\xi,\,l}$ $(l = 1,\ \ldots,\ L)$ *zerlegt werden, d. h.*

$$\mathbb{L}_L^{\mathbb{C},\,\xi} = \bigcup_{l=1}^{L}\ \mathbb{L}_L^{\mathbb{C},\,\xi,\,l}.$$

Dabei ist $\mathbb{L}_L^{\mathbb{C},\,\xi,\,l}$ *von folgender Gestalt:*

$$\mathbb{L}_L^{\mathbb{C},\,\xi,\,l} = \Big\{ [\Phi]_\sim \,\in\, \mathbb{L}_L^{\mathbb{C}} :\; \dim\big(\Phi\ \mathbb{C}^L \cap \Psi_\xi\ \mathbb{C}^L\big) = l \Big\}.$$

Natürlich kann man auch allgemeinere singuläre Bott-Maslov-Zykel konstruieren, die nicht auf Lagrange-Ebenen der Form Ψ_ξ basieren (vgl. [Ar85]), aber da die obige Definition für die Zwecke dieser Arbeit genügt, werde ich mich auf diese beschränken. Weiterhin ist es aufschlussreich, die Schnittbedingung mit Hilfe der Wronski-Matrix auszudrücken.

Definition 6.2 (Wronski-Matrix)
Die Wronski-Matrix $\mathcal{W}$ *von zweier Repräsentanten* Φ *und* Ψ *von Lagrange-Ebenen* $[\Phi]_\sim$ *und* $[\Psi]_\sim$ *in* $\mathbb{L}_L^{\mathbb{C}}$ *hat die Darstellung*

$$\mathcal{W}(\Phi,\ \Psi) = \Phi^*\ \mathcal{J}\ \Psi.$$

[4] Auch dieser Zykel wird in der einschlägigen Literatur oft als „Maslov-Zykel“ bezeichnet – doch da auch diese Konstruktion schon bei Bott (in [Bo56]) auftaucht, sollte man gerechterweise zumindest von einem „Bott-Maslov-Zykel“ sprechen.

Bemerkung 6.1 *Hilfreich für ein intuitives Verständnis des folgenden Sachverhaltes ist es, sich vorweg die* Wronski-Determinante[5] *vor Augen zu führen:*

$$\begin{aligned}\det\left(\mathcal{W}(\Phi,\,\Psi)\right) &= \det\left(\Phi^*\,\mathcal{J}\,\Psi\right)\\ &= \det\left(\begin{pmatrix}a^* & b^*\end{pmatrix}\begin{pmatrix}0 & -\mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}c\\ d\end{pmatrix}\right)\\ &= \det\left(b^*c - a^*d\right).\end{aligned}$$

Lemma 6.1 *Für Repräsentanten zweier Lagrange-Ebenen* $\Phi \in [\Phi]_\sim$ *und* $\Psi \in [\Psi]_\sim$ *(mit* $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ *und* $[\Psi]_\sim \in \mathbb{L}_L^{\mathbb{C}}$*) gilt:*

$$\dim\left(\Phi\ \mathbb{C}^L\ \cap\ \Psi\ \mathbb{C}^L\right) = \dim\Big(\ker\ \mathcal{W}\big(\Phi,\,\Psi\big)\Big).$$

Beweis: Es sei zunächst eine Bemerkung zur Notation gestattet: Man setze $\Phi = \begin{pmatrix}a\\ b\end{pmatrix}$ und $\Psi = \begin{pmatrix}\widehat{a}\\ \widehat{b}\end{pmatrix}$. Darüber hinaus sei daran erinnert, dass Φ genau dann eine Lagrange-Ebene repräsentiert, wenn $\Phi^*\mathcal{J}\Phi = 0$ bzw. $a^*b = b^*a$ ist.

Nun sei φ_i eine Basis von $\Phi\,\mathbb{C}^L \cap \Psi\,\mathbb{C}^L$ ($i = 1,\ \ldots,\ l,\ l \leq 2L$). Dann exisitieren $\vartheta^i_{1,\,2} \in \mathbb{C}^L,\ \vartheta^i_{1,\,2} \neq 0$, für die gilt:

$$\varphi_i = \Phi\vartheta^i_1 = \Psi\vartheta^i_2 \qquad \text{für alle } i = 1,\ \ldots,\ l.$$

Daraus folgt:

$$\begin{aligned}0 &= (\vartheta^i_1)^*\quad \underbrace{\Phi^*\mathcal{J}\Phi}_{=0,\ \text{Lemma 3.3}}\quad \vartheta^i_1\\ &= (\vartheta^i_1)^*\ \Phi^*\mathcal{J}\Psi\ \vartheta^i_2.\end{aligned}$$

Daher ist $\dim\big(\Phi\ \mathbb{C}^L \cap \Psi\ \mathbb{C}^L\big) \leq \dim\Big(\ker\big(\mathcal{W}(\Phi,\ \Psi)\big)\Big)$.

---

[5]Bekannt und hilfreich zur Lösung von Differentialgleichungen höherer Ordnung ist die *Wronski-Determinante* auf einem $2L$-dimensionalen Lösungsraum:

$$\det\big(\mathcal{W}(Y_1, Y_2)\big) = \sum_{i=1}^{L}(y_1^i)'\,y_2^i - y_1^i(y_2^i)'.$$

Dies ist eine nichtdegenerierte, schiefsymmetrische Form auf $\mathbb{C}^{2L\times 2L}$. Damit ist sie nach A.2 isomorph zur symplektischen Standardform ω_0 auf $\mathbb{C}^{2L\times 2L}$.

Benannt sind sowohl die Wronski-Determinante als auch die Wronski-Matrix nach dem polnischen Mathematiker *Graf Hoené Wronski* (1778–1853), der diese im Jahr 1821 einführte.

Jetzt nimmt man an, dass $\dim\Big(\ker\big(\mathcal{W}(\Phi,\ \Psi)\big)\Big) > \dim\big(\Phi\,\mathbb{C}^L \cap \Psi\,\mathbb{C}^L\big)$ ist. Dazu setze man $\dim\big(\Phi\,\mathbb{C}^L \cap \Psi\,\mathbb{C}^L\big) = l$, ohne Einschränkung sei $\dim\Big(\ker\big(\mathcal{W}(\Phi,\ \Psi)\big)\Big) = l+1$. Dann gibt es eine Basis $\{v_1,\ \ldots,\ v_{l+1}\}$, $v_i \in \mathbb{C}^L$ dieses Kerns und für jedes v_i $(i = 1,\ \ldots,\ l+1)$ gilt: $v_i \in \ker\big(\mathcal{W}(\Phi,\ \Psi)\big)$, d. h. $\big(\mathcal{W}(\Phi,\ \Psi)\big)\, v_i = 0$ bzw. $\Phi^* \mathcal{J} \Psi\, v_i = 0$. Nutzt man wiederum die Schreibweise in Blockmatrixform, ergibt sich: $b^*\, c\, v_i = a^*\, d\, v_i$. Diese Bedingung ist nach Folgerung 3.1 erfüllt, wenn $\widehat{a}\, v_i = a\, v_i$ und $\widehat{b}\, v_i = b\, v_i$ ist. Setzt man dies wieder zusammen, ergibt sich: $\begin{pmatrix} \widehat{a}\, v_i \\ \widehat{b}\, v_i \end{pmatrix} = \begin{pmatrix} a\, v_i \\ b\, v_i \end{pmatrix}$, d. h. v_i liegt im Schnitt $\Phi\,\mathbb{C}^L \cap \Psi\,\mathbb{C}^L$. Die v_i sind aber, da sie eine Basis bilden, linear unabhängig voneinander. Das ist ein Widerspruch dazu, dass die Dimension des Schnittes, $\dim\big(\Phi\,\mathbb{C}^L \cap \Psi\,\mathbb{C}^L\big)$, gleich l ist. Damit ist der Beweis ist abgeschlossen. □

Bemerkung 6.2
Aus Bemerkung 6.1 und dem Beweis zu Lemma 6.1 folgt, dass gilt:

$$\Phi\,\mathbb{C}^L \ \cap\ \Psi\,\mathbb{C}^L \neq \{\overrightarrow{0}\} \ \Leftrightarrow\ \det\big(\mathcal{W}(\Phi,\ \Psi)\big) = 0.$$

Aus dem Beweis zu Lemma 6.1 geht ebenso hervor, dass die Dimension des Kerns, $\dim\Big(\ker\big(\mathcal{W}(\Phi,\ \Psi)\big)\Big)$, nicht *von der Wahl der Basen der Lagrange-Ebenen* $[\Phi]_\sim$ *und* $[\Psi]_\sim$ *abhängt – obwohl der konkrete Wert* $\mathcal{W}(\Phi,\ \Psi)$ *im allgemeinen sehr wohl von der Wahl der Basen abhängig ist.*

Nach Definition 3.3 gilt, dass aus $[\Phi]_\sim = [\Psi]_\sim$ *folgt:*

$$\mathcal{W}(\Phi,\ \Psi) = \Phi^* \mathcal{J} \Psi = \Phi^* \mathcal{J} \Phi\, c = 0.$$

Umgekehrt gilt im Falle, dass $\mathcal{W}(\Phi,\ \Psi) = 0$ *ist (mit den Bezeichnungen aus dem Beweis von Lemma 6.1):*

$$0 = \Phi^* \mathcal{J} \Psi = b^* \widehat{a} - a^* \widehat{b}.$$

Dies ist aber genau dann erfüllt, falls ein $c \in Gl(L, \mathbb{C})$ *exisitiert, so dass* $b^*ac - a^*bc$ *ist. Insgesamt ergibt sich:*

$$\mathcal{W}(\Phi,\ \Psi) = 0 \Leftrightarrow [\Phi]_\sim = [\Psi]_\sim.$$

Das nächste Ziel dieses Abschnitts ist es, einen Schnittindex zu definieren. Weithin bekannt ist das Resultat von Arnol'd, der in [Ar67] gezeigt hat, dass der Bott-Maslov-Zykel aus Definition 6.1 auf $\mathbb{L}_L^{\mathbb{R},\xi}$ einen koorientierten Zykel der Kodimension 1 definiert. Damit ist die Zweiseitigkeit von $\mathbb{L}_L^{\mathbb{R},\xi}$ etabliert und es ist möglich, eine Richtung zu definieren. Schon vor Arnol'd zeigte Bott in [Bo56], dass man auf $\mathbb{L}_L^{\mathbb{C},\,\xi,\,l}$ eine Richtung definieren kann, mit deren Hilfe man die Wohldefiniertheit der folgenden Schnittindizes sicherstellen kann. Diese Möglichkeit eröffnet auch der folgende Satz.

Satz 6.1 *Es gilt:*

$$\Pi\left(\mathbb{L}_L^{\mathbb{C},\,\xi,\,l}\right) = \partial_L^{\xi,\,l}\mathbb{D}_L^{\mathbb{C}}.$$

Dabei ist

$$\partial_L^{\xi,\,l}\mathbb{D}_L^{\mathbb{C}} = \left\{U \;\in\; \partial_L\,\mathbb{D}_L^{\mathbb{C}} :\; \mathrm{Rang}\left(\Pi([\Psi_\xi]_\sim - U\right) = L - l\right\}.$$

Setzt man weiterhin $\partial_L^{\xi}\,\mathbb{D}_L^{\mathbb{C}} = \bigcup_{l=1}^{L}\partial_L^{\xi,\,l}\mathbb{D}_L^{\mathbb{C}}$*, so folgt:*

$$\Pi\left(\mathbb{L}_L^{\mathbb{C},\xi}\right) = \partial_L^{\xi}\mathbb{D}_L^{\mathbb{C}}.$$

Außerdem gilt für jedes $\xi \;\in\; \mathrm{Her}(L,\mathbb{C})$*:* $\mathbb{1} \;\notin\; \partial_L^{\xi}\mathbb{D}_L^{\mathbb{C}}$.

Beweis: Sei $U \;\in\; \partial_L\mathbb{D}_L^{\mathbb{C}}$. Man setze – wie gewohnt – den Repräsentanten Φ von $[\Phi]_\sim$ als $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$. Hier sei $a = \frac{1}{2}(U + \mathbb{1})$ und $b = \frac{i}{2}(U - \mathbb{1})$ wie in (4.3). Damit ist

$$\begin{aligned}
\Pi([\Phi]_\sim) &= (a - ib)(a + ib)^{-1} \\
&= \left(\frac{1}{2}(U + \mathbb{1}) - i\frac{i}{2}(U - \mathbb{1})\right)\left(\frac{1}{2}(U + \mathbb{1}) + i\frac{i}{2}(U - \mathbb{1})\right)^{-1} \\
&= \left(\frac{1}{2}U + \frac{1}{2}\mathbb{1} + \frac{1}{2}U - \frac{1}{2}\mathbb{1}\right)\left(\frac{1}{2}U + \frac{1}{2}\mathbb{1} - \frac{1}{2}U + \frac{1}{2}\mathbb{1}\right)^{-1} \\
&= U.
\end{aligned}$$

Für die folgenden Rechnungen ist es hilfreich, $\Pi(\Psi_\xi)$ explizit anzugeben. Vorweg sei betont, dass $(\xi + i\mathbb{1})$ invertierbar ist, weil die Eigenwerte der hermiteschen Matrix ξ allesamt reell sind. Daher sind alle Eigenwerte von $(\xi + i\mathbb{1})$ ungleich 0, was gleichbedeutend mit der Invertierbarkeit ist. Somit gilt:

$$\Pi(\Psi_\xi) = \Pi\begin{pmatrix} \xi \\ \mathbb{1} \end{pmatrix} = (\xi - i\mathbb{1})(\xi + i\mathbb{1})^{-1}.$$

Daraus folgt:

$$\begin{aligned}
\mathcal{W}(\Psi_\xi,\ \Phi) &= \begin{pmatrix}\xi\\ \mathbb{1}\end{pmatrix}^* \begin{pmatrix}0 & -\mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \begin{pmatrix}\mathbb{1} & & -\xi^*\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}\\
&= a - \xi^* b \overset{\xi^*=\xi}{=} a - \xi b\\
&= \frac{1}{2}(U+\mathbb{1}) - \frac{i\xi}{2}(U-\mathbb{1}) = \frac{1}{2}\big(U + \mathbb{1} - i\xi(U-\mathbb{1})\big)\\
&= \frac{1}{2}\big((\mathbb{1}-i\xi)U + \mathbb{1} + i\xi\big)
\end{aligned}$$

Außerdem gilt:

$$\begin{aligned}
\mathbb{1} + i\xi &= \frac{(\xi + i\mathbb{1})(\mathbb{1}+i\xi)}{\xi + i\mathbb{1}} = \frac{(\mathbb{1}-i\xi)\, i^2\, (\xi - i\mathbb{1})}{\xi + i\mathbb{1}}\\
&= (\mathbb{1}-i\xi)\left(-\frac{\xi - i\mathbb{1}}{\xi + i\mathbb{1}}\right) = -(\mathbb{1}-i\xi)\ \Pi([\Psi_\xi]_\sim).
\end{aligned}$$

Daraus folgt (man beachte dabei, dass $U = \Pi\,([\Phi]_\sim)$ ist):

$$\frac{1}{2}\big((\mathbb{1}-i\xi)U + \mathbb{1} + i\xi\big) = \frac{1}{2}(\mathbb{1}-i\xi)\big(U - \Pi([\Psi_\xi]_\sim)\big).$$

Insgesamt ergibt sich:

$$\mathcal{W}(\Psi_\xi,\ \Phi) = \frac{1}{2}(\mathbb{1}-i\xi)\big(U - \Pi([\Psi_\xi]_\sim)\big).$$

Da $\mathbb{1} - i\xi$ aus den gleichen Gründen wie $\mathbb{1} + i\xi$ invertierbar ist, erhält man aus den bisherigen Überlegungen und unter Beachtung von Lemma 6.1 folgendes Resultat:

$$\begin{aligned}
\dim\big(\Phi\ \mathbb{C}^L \cap \Psi_\xi\ \mathbb{C}^L\big) &= \dim\Big(\ker\big(\mathcal{W}(\Psi_\xi\ \Phi)\big)\Big)\\
&= \dim\Big(\ker\big(\Pi([\Phi]_\sim) - \Pi([\Psi_\xi]_\sim)\big)\Big).
\end{aligned}$$

Da der Zykel $\mathbb{L}_L^{\mathbb{C},\,\xi,\,l}$ durch

$$\mathbb{L}_L^{\mathbb{C},\,\xi,\,l} = \Big\{[\Phi]_\sim\ \in\ \mathbb{L}_L^{\mathbb{C}} :\ \dim\big(\Phi\ \mathbb{C}^L \cap \Psi_\xi\ \mathbb{C}^L\big) = l\Big\}$$

definiert ist, folgt für $\Phi\ \in\ \mathbb{L}_L^{\mathbb{C},\,\xi,\,l}$:

$$\dim\Big(\ker\big(\Pi([\Psi_\xi]_\sim) - \Pi([\Phi]_\sim)\big)\Big) = \dim\Big(\ker\big(\Pi([\Psi_\xi]_\sim) - U\big)\Big) = l.$$

Nimmt man nun an, dass $\mathbb{1} \in \partial_L^\xi \mathbb{D}_L^\mathbb{C}$ ist, so bedeutet dies, dass ein $l \in \{1, \ldots, L\}$ dergestalt existiert, dass gilt

$$\operatorname{Rang}\left(\Pi\left([\Psi_\xi]_\sim\right) - \mathbb{1}\right) = L - l.$$

Da nach Definition gilt:

$$\Phi = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\mathbb{1} + \mathbb{1}) \\ \frac{i}{2}(\mathbb{1} - \mathbb{1}) \end{pmatrix} = \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix},$$

folgt:

$$\mathcal{W}(\Psi_\xi,\ \Phi) = \mathbb{1} - \xi 0 = \mathbb{1}.$$

Daher ist $\operatorname{Rang}(\mathcal{W}(\Psi_\xi,\ \Phi)) = L$, die Dimension des Kernes $\dim(\ker(\mathcal{W}(\Psi_\xi,\ \Phi)))$ ist dementsprechend gleich 0. Dies ist aber ein Widerspruch dazu, dass $l \in \{1, \ldots, L\}$ ist. Also ist $\mathbb{1} \notin \partial_L^\xi \mathbb{D}_L^\mathbb{C}$.

Damit ist der Beweis des Satzes komplett. □

Schlussfolgerung

Ich werde mich im weiteren Verlauf dieser Arbeit nur auf den singulären Zykel stützen, der zu $\xi = -\cot\left(\frac{\varphi}{2}\right)\mathbb{1},\ \varphi \in (0,\ 2\pi)$ assoziiert ist. Diesen singulären Zykel bezeichne ich mit $\mathbb{L}_L^\varphi$. Natürlich ist $\mathbb{L}_L^\varphi \subset \mathbb{L}_L^\mathbb{C}$, und die für den allgemeinen Fall bewiesenen Aussagen verlieren ihre Gültigkeit nicht. Speziell hat man auch in diesem Falle wieder die disjunkte Zerlegung $\mathbb{L}_L^\varphi = \bigcup_{l=1}^L \mathbb{L}^{\varphi,\,l}$ zur Verfügung. Das Bild dieses Zykels unter Π nenne ich $\partial_L^\varphi\, \mathbb{D}_L^\mathbb{C} = \bigcup_{l=1}^L \partial_L^{\varphi,\,l}\, \mathbb{D}_L^\mathbb{C}$. Mit diesen Konventionen gilt:

$$\begin{aligned}
\Pi(\Psi_\xi) &= \Pi(\Psi_{-\cot\left(\frac{\varphi}{2}\right)\mathbb{1}}) \\
&= \left(-\cot\left(\frac{\varphi}{2}\right)\mathbb{1} - i\mathbb{1}\right)\left(-\cot\left(\frac{\varphi}{2}\right)\mathbb{1} + i\mathbb{1}\right)^{-1} \\
&= \left(\frac{\cos(-\varphi/2)\mathbb{1} - i\ \sin(-\varphi/2)\mathbb{1}}{\sin(-\varphi/2)\mathbb{1}}\right)\left(\frac{\sin(-\varphi/2)\mathbb{1}}{\cos(-\varphi/2)\mathbb{1} + i\ \sin(-\varphi/2)\mathbb{1}}\right) \\
&= \frac{\cos(-\varphi/2)\mathbb{1} + i\ \sin(\varphi/2)\mathbb{1}}{\cos(-\varphi/2)\mathbb{1} + i\ \sin(-\varphi/2)\mathbb{1}} \\
&= \frac{\cos(\varphi/2)\mathbb{1} + i\ \sin(\varphi/2)\mathbb{1}}{\cos(-\varphi/2)\mathbb{1} + i\ \sin(-\varphi/2)\mathbb{1}} \\
&= e^{i\varphi/2}\,(e^{-i\varphi/2})^{-1}\ \mathbb{1} \\
&= e^{i\varphi}\ \mathbb{1}.
\end{aligned}$$

Nach (5.3) ist $\mathbb{D}_L^{\mathbb{C}} = \mathfrak{U}(L)$. Zusammen mit Satz 6.1 folgt damit diese äußerst wichtige Aussage:

$$\begin{aligned}\Pi(\mathbb{L}_L^{\varphi,\,l}) &= \partial_L^{\varphi,\,l}\,\mathbb{D}_L^{\mathbb{C}} \\ &= \left\{U \in \partial_L \mathbb{D}_L^{\mathbb{C}} :\ \text{Rang}\left(e^{i\varphi}\,\mathbb{1} - U\right) = L - l\right\} \\ &= \left\{U \in \mathfrak{U}(L) :\ e^{i\varphi} \text{ ist Eigenwert von } U \text{ mit Multiplizität } l\right\}.\end{aligned} \tag{6.2}$$

6.2 Der Schnittindex (Bott-Maslov-Index)

Vorweg sei darauf erinnert, dass $\mathbb{L}_L^{\varphi} = \bigcup_{l=1}^{L} \mathbb{L}^{\varphi,l}$ ist. Jedes dieser $\mathbb{L}^{\varphi,l}$ ist eine algebraische Untermannigfaltigkeit von $\mathbb{L}_L^{\mathbb{C}}$ mit Kodimension $\frac{l(l+1)}{2}$. Speziell folgt daraus, dass die Kodimension von $\mathbb{L}^{\varphi,1}$ gleich $1(1+1)/2 = 1$ ist. All diese Resultate gehen auf Arnol'd ([Ar67]) zurück. Ebenfalls in [Ar67] hat Arnol'd gezeigt, dass $\mathbb{L}_L^{\mathbb{R},\xi,1}$ beidseitig in $\mathbb{L}_L^{\mathbb{R}}$ eingebettet ist. Das bedeutet, dass ein stetiges Vektorfeld existiert, das tangential zu $\mathbb{L}_L^{\mathbb{R}}$ und transversal zu $\mathbb{L}_L^{\mathbb{R},\xi}$ ist. Damit ist es möglich, einen (gewichteten) Index von Schnitten zu definieren, da man sich auf eine positive und eine negative Seite von $\mathbb{L}_L^{\mathbb{R},\xi}$ beziehen kann. Das gilt natürlich auch für alle hier gewählte ξ. Bott zeigte schon früher (in [Bo56]), dass ein solches Resultat auch für $\mathbb{L}_L^{\mathbb{C},\xi}$ gilt.[6] Außerdem sei erwähnt, dass es weitere Ansätze gibt, um den Bott-Maslov-Index zu bestimmen.[7] In [FJN03] wird gezeigt, dass all diese Ansätze äquivalent sind, daher genügt es, den hier gewählten Bott-Maslov-Index zu studieren.

Nun die Ausführungen en détail: Sei $\Gamma = \left([\Phi^E]_\sim\right)_{E \in [E_0,E_1)}$ ein abgeschlossener, stetiger Pfad in $\mathbb{L}_L^{\mathbb{C}}$, für den die Menge der Schnitte mit $\mathbb{L}_L^{\varphi}$, d. h. $\left\{E \in [E_0, E_1) :\ \Gamma(E) \in \mathbb{L}_L^{\varphi}\right\}$, endlich ist.

Jetzt werden diese Schnittstellen genauer analysiert: Man wähle $E \in \mathbb{R}$ so, dass $\Gamma(E) \in \mathbb{L}_L^{\varphi,\,l}$ ist. Außerdem seien $\theta_1(\widetilde{E}),\ \ldots,\ \theta_l(\widetilde{E})$ die Eigenphasen der unitären Matrix[8] $\Pi\big(\Gamma(\widetilde{E})\big)$, die an der Stelle $\widetilde{E} = E$ allesamt gleich φ sind. Im nächsten Schritt wählt man $\epsilon,\ \delta > 0$ so, dass für $\widetilde{E} \in [E-\epsilon,\ E+\epsilon]$ die im ersten Schritt gewählten $\theta_k(\widetilde{E})$ $(k = 1,\ \ldots,\ l)$ in $[\varphi-\delta,\ \varphi+\delta]$ sind. Dabei ist wichtig, dass

[6] Der hier betrachtete Fall ist allerdings etwas allgemeiner, da Bott und Arnol'd die Resultate nur für Pfade in sogenannter generischer Position bewiesen haben.

[7] Diese Ansätze stützen sich auf das sogenannte *Argument* symplektischer Matrizen. Dabei versteht man unter diesem Argument den Winkel, der sich ergibt, wenn eine symplektische Matrix über $\mathbb{S}^1$ projiziert wird. Die Rotationszahl wird dann bestimmt über das mittlere Inkrement dieses Arguments (vgl. [NNO98]).

[8] Da alle Eigenwerte einer unitären Matrix vom Betrag 1 sind, bewegen sich die Eigenphasen auf dem Einheitskreis.

ϵ so bestimmt wurde, dass für $\widetilde{E} \neq E$ außer den oben gewählten keine anderen Eigenphasen in $[\varphi - \delta,\ \varphi + \delta]$ sind. Weiterhin sind die Eigenphasen bei φ nicht konstant und schneiden φ nur genau einmal auf dem Intervall $[E - \epsilon,\ E + \epsilon]$. Anders ausgedrückt: Für $\widetilde{E} \in [E - \epsilon,\ E + \epsilon]$ gilt: Falls $\widetilde{E} \neq E$ ist, ist $\theta_k(\widetilde{E}) \neq \varphi$.

Definition 6.3 (Signatur)
Sei n_- bzw. n_+ die Anzahl der l Eigenphasen, die vor *(n_-) bzw.* nach *(n_+) dem Schnitt kleiner als φ sind. Analog seien p_- und p_+ die Anzahl der l Eigenphasen, die* vor *(p_-) bzw.* nach *(p_+) dem Schnitt größer als φ sind. Dann ist die* Signatur von $\Gamma(E)$ *definiert durch:*

$$\operatorname{sgn}\bigl(\Gamma(E)\bigr) = \frac{1}{2}\bigl(p_+ - n_+ - p_- + n_-\bigr).$$

Nach Definition von n_-, n_+ und p_-, p_+ gilt: $n_- + n_+ = l$ und $p_- + p_+ = l$. Das ist äquivalent zu: $l - n_+ = n_-$ und $l - p_- = p_+$. Setzt man dies ein, erhält man:

$$\operatorname{sgn}\ \bigl(\Gamma(E)\bigr) = l - n_+ - p_-. \tag{6.3}$$

Offensichtlich nimmt die Signatur nur Werte in $\mathbb{Z}$ an. Da $0 \leq n_+,\ p_- \leq l$ ist, folgt aus dieser Definition sofort, dass $-l \leq \operatorname{sgn}\ \bigl(\Gamma(E)\bigr) \leq l$ ist. Wichtig ist außerdem, dass diese Signatur unabhängig von der gewählten Reihenfolge der Eigenphasen am Schnittpunkt ist.

Wie kann man diesen Ausdruck geometrisch interpretieren?[9] Erfreulicherweise kann man sich die Signatur sehr leicht veranschaulichen, denn $\operatorname{sgn}\bigl(\Gamma(E)\bigr)$ entspricht genau der Anzahl aller Eigenphasen, die φ bei E treffen, abzüglich der Eigenphasen die nach dem Schnitt kleiner als φ sind und abzüglich jener, die vor dem Schnitt größer als φ waren. Somit entspricht $\operatorname{sgn}\bigl(\Gamma(E)\bigr)$ der Anzahl der Eigenphasen, die φ gegen den Uhrzeigersinnn durchquert haben.

Unter bestimmten Voraussetzungen kann man $\operatorname{sgn}\bigl(\Gamma(E)\bigr)$ auch mit Hilfe der partiellen Ableitungen der Eigenphasen erklären. Eine erste Voraussetzung ist die Differenzierbarkeit der Eigenphasen. Darüber hinaus ist es notwendig, auch $\partial_E\ \theta_k(E) \neq 0$ für $k = 1,\ \ldots,\ l$ zu fordern. Damit wird ausgeschlossen, dass φ nur ein Tangentialpunkt einer der Phasen ist, und es ist sichergestellt, dass alle Phasen φ tatsächlich „durchstoßen“. Damit ist offensichtlich, dass das Vorzeichen der partiellen Ableitung $\partial_E\ \theta_k(E)$ die Richtung angibt, in der die Eigenphase θ_k die Schnittstelle φ durchläuft. Ist $\partial_E\ \theta_k(E) > 0$, so durchquert θ_k die Stelle φ

[9] Für eine ausführlichere Darstellung der geometrischen Eigenschaften siehe [CLR90]. Dort finden sich auch einige Skizzen.

gegen den Uhrzeigersinn und für $\partial_E \theta_k(E) < 0$ durchquert θ_k φ im Uhrzeigersinn. Daraus folgt:

$$\operatorname{sgn}\bigl(\Gamma(E)\bigr) = \sum_{k=1}^{l} \operatorname{sgn}(\partial_E \theta_k(E)).$$

Diese Vorarbeiten erlauben es nun, den Schnittindex eines Pfades Γ zu definieren. Die hier gewählte, sehr klare und einfache Definition ist übernommen von [SB06].

Definition 6.4 *Der Bott-Maslov-Index eines Pfades Γ hinsichtlich eines singulären Zykels $\mathbb{L}_L^{\varphi}$ ist definiert durch*

$$\operatorname{ind}\ (\Gamma,\ \mathbb{L}_L^{\varphi}) = \sum_{\Gamma(E)\,\in\,\mathbb{L}_L^{\varphi}} \operatorname{sgn}\ \bigl(\Gamma(E)\bigr). \tag{6.4}$$

Um den Bott-Maslov-Index explizit zu berechnen, ist es hilfreich, den Index mit Hilfe von *Windungszahlen* auszudrücken. Es sei kurz daran erinnert, dass eine Windungszahl eine topologische Invariante ist, die – grob gesagt – für eine geschlossene Kurve Γ und einen Punkt z_0 misst, wie oft der Punkt z_0 entgegen dem Uhrzeigersinn umlaufen wird, wenn man dem Verlauf der Kurve folgt. Um einzusehen, dass der Bott-Maslov-Index tatsächlich einer Windungszahl entspricht, definiert man zunächst die Kurve, deren Verlauf zu folgen ist. Dazu seien die Abbildungen $[E_0,\ E_1] \ni E \mapsto \theta_l^E$ die stetigen Pfade der Eigenphasen von $\Pi\bigl(\Gamma(E)\bigr)$, die wie in Definition 6.3 definiert sind. Dabei ist die Reihenfolge am Schnittpunkt zufällig gewählt. Da diese Pfade geschlossen sind, kann jedem von ihnen eine Windungszahl zugeordnet werden.

Nun analysiert man, wie sich ein Umlauf einer Eigenphase auf diesem Pfad einerseits auf den Bott-Maslov-Index und andererseits auf die Windungszahl auswirkt. Da die Pfade der Eigenphasen nach Voraussetzung geschlossen sind, *muss* jeder der vorgegeben Punkte $\varphi \in \mathbb{L}_L^{\varphi}$ bei einem kompletten Umlauf $n+1$-mal durchquert werden ($n \in \mathbb{N}_0$). Nach Definition der Signatur $\operatorname{sgn}\bigl(\Gamma(E)\bigr)$ heben sich zwei Durchquerungen in unterschiedliche Richtungen in der Summe gerade auf. Die Signatur $\operatorname{sgn}\bigl(\Gamma(E)\bigr)$ wird sich beim kompletten Umlauf einer Eigenphase um $+1$ oder -1 verändern: Falls θ_k einen Punkt z_0 gegen den Uhrzeigersinn umläuft[10], wird φ einmal mehr gegen den Uhrzeigersinn als im Uhrzeigersinn durchstoßen. Damit ist die Signatur eines Umlaufes von θ_k gleich 1. Analog gilt, dass die Signatur für einen Umlauf im Uhrzeigersinn gleich -1 ist. Das entspricht aber genau

[10]Dabei nimmt man an, dass z_0 im Inneren des von den Eigenphasen beschriebenen Kreises liegt.

der Definition der Windungszahl. Damit kann man nun den Bott-Maslov-Index als die Summe dieser Windungszahlen ausdrücken (vgl. z. B. auch [Ro92]).

$$\operatorname{ind}(\Gamma,\ \mathbb{L}_L^{\varphi}) = \sum_{l=1}^{L} \operatorname{Wind}\big(E\ \in\ [E_0, E_1) \mapsto \theta_l^E\big)\,. \tag{6.5}$$

Weitere alternative Formulierungen des Bott-Maslov-Indexes findet, wie schon erwähnt, beispielsweise in [FJN03].

6.3 Der Kozyklus nach Arnol'd

Die nun folgenden Resultate sind essentiell für die explizite Berechnung des Bott-Maslov-Indexes.

Bereits Arnol'd hat (in [Ar67]) gezeigt, dass $H^1(\mathbb{L}_L^{\mathbb{R}},\ \mathbb{Z}) \simeq \mathbb{Z}$ ist. Aus Theorem 1 ergibt sich die allgemeinere Folgerung $H^1(\mathbb{L}_L^{\mathbb{C}}, \mathbb{Z}) \simeq \mathbb{Z}$. Um diesen Zusammenhang zu sehen, wähle man den Generator ω der de-Rham-Gruppe wie folgt: Sei $\Gamma = \big([\Phi^E]_\sim\big)_{E \in [E_0, E_1)}$ ein geschlossener, stetiger Pfad in $\mathbb{L}_L^{\mathbb{C}}$. Mit Hilfe dieses Pfades konstruiert man die Abbildung $[E_0,\ E_1) \ni E \mapsto \det\big(\Pi([\Phi^E]_\sim)\big)$. Diese Abbildung ist wieder ein geschlossener, stetiger Pfad. Nach Definition von Π als unitärer Matrix (und natürlich von nach der Definition der Determinante) ist dieser neue Pfad, $\det\big(\Pi([\Phi^E]_\sim\big)$, in $\mathbb{S}^1$. Seine Windungszahl basiert auf der Paarung der de Rham-Klasse des Generators ω und der Homotopieklasse des Pfades Γ:

$$\langle\, \omega \mid \Gamma \,\rangle = \operatorname{Wind}\Big(E\ \in\ [E_0,\ E_1) \mapsto \det\big(\Pi([\Phi^E]_\sim)\big)\Big). \tag{6.6}$$

Da jeder stetige Pfad in $\mathbb{L}_L^{\mathbb{C}}$ durch einen differenzierbaren Pfad approximiert werden kann, setzt man von nun an o. B. d. A. den Pfad Γ als differenzierbar voraus. Somit kann man die Paarung aus (6.6) berechnen durch:

$$\langle\, \omega \mid \Gamma \,\rangle = \int_{E_0}^{E_1} \frac{dE}{2\pi}\ \Im\ \partial_E \log\Big(\det\big(\Pi([\Phi^E]_\sim)\big)\Big). \tag{6.7}$$

Der nächste Satz ist Arnol'ds (vgl. [Ar67]) Verdienst und von herausragender Bedeutung für das Verständnis des oben definierten Kozykels, denn er stellt den Zusammenhang zwischen der Paarung $\langle\, \omega \mid \Gamma \,\rangle$ und dem Bott-Maslov-Index her.

Satz 6.2 *Sei Γ ein geschlossener, differenzierbarer Pfad, der nur endlich viele Schnitte mit einem singulären Zykel $\mathbb{L}_L^{\varphi}$ hat. Dann ist*

$$\operatorname{ind}(\Gamma,\ \mathbb{L}_L^{\varphi}) = \langle\, \omega \mid \Gamma \,\rangle.$$

Beweis (vgl. [SB06]): Sei $U = U(E) = \Pi\big(\Gamma(E)\big)$. Wesentlich für den Beweis ist, dass jede symmetrische unitäre Matrix U durch eine unitäre, differenzierbare Matrix $M \in \mathfrak{U}(L)$ diagonalisiert werden kann. D. h. $U = M^*\, D\, M$, wobei $D = \operatorname{diag}\left(e^{i\theta_1^E},\ \dots,\ e^{i\theta_L^E}\right)$ mit $\theta_l^E \in [0,\ 2\pi)$ für $l = 1,\ \dots,\ L$ ist.[11]

Da $M = M^E$ in $\mathfrak{U}(L)$ ist, gilt: $(M^E)^* = (M^E)^{-1}$. Daher ist

$$\begin{aligned}
M^E(\partial_E\, M^E)^* &= M^E(\partial_E\,(M^E)^{-1}) \\
&= M^E\big(-(M^E)^{-1}\,\partial_E\, M^E\,(M^E)^{-1}\big) \\
&= -(\partial_E M^E)(M^E)^{-1} \\
&= -(\partial_E M^E)(M^E)^*
\end{aligned}$$

und es folgt[12]:

$$\begin{aligned}
& \Im\left(\partial_E\ \log\Big(\det\big(U(E)\big)\Big)\right) \\
=\ & \Im\left(\partial_E\ \log\Big(\det\big((M^E)^*\, D^E\, M^E\big)\Big)\right) \\
=\ & \frac{1}{i}\ \mathrm{Spur}\ \Big(\ \big((M^E)^*\, D^E\, M^E\big)^*\ \big(\partial_E\ ((M^E)^*\, D^E\, M^E)\big)\Big) \\
=\ & \frac{1}{i}\ \mathrm{Spur}\ \Big(\ (M^E)^*(D^E)^* M^E\big((\partial_E\,(M^E)^*)\, D^E\, M^E \\
& \qquad +(M^E)^*\ (\partial_E\ D^E)\ M^E + (M^E)^*\ D^E\ (\partial_E M^E)\big)\Big) \\
=\ & \frac{1}{i}\ \mathrm{Spur}\ \Big((M^E)^*(D^E)^*\,\underbrace{M^E(M^E)^*}_{=\mathbb{1}}\,(\partial_E\ D^E)\ M^E\Big) \\
=\ & \frac{1}{i}\ \mathrm{Spur}\ \Big((M^E)^*(D^E)^*(\partial_E\ D^E)\ M^E\Big)
\end{aligned}$$

[11] Da U von E abhängt, sind natürlich auch die Matrizen M und D von E abhängig. Falls auf diesen Zusammenhang hingewiesen werden soll, verwende ich die Notation M^E bzw. D^E.

[12] Dabei beachte man, dass die Spur invariant unter zyklischen Vertauschungen ist, d. h. Spur (ABC) = Spur (CAB) = Spur (CBA).

$$
\begin{aligned}
&= \frac{1}{i}\,\mathrm{Spur}\,\Big(\big(D^E\big)^* \partial_E D^E\Big)\\
&= \frac{1}{i}\,\mathrm{Spur}\,\Big(\mathrm{diag}\big(e^{-i\theta_1^E},\ \ldots,\ e^{-i\theta_1^E}\big)\ \ \mathrm{diag}\big(i\,\partial_E\,\theta_L^E\,e^{i\,\theta_1^E},\ \ldots,\ i\,\partial_E\,\theta_L^E\,e^{i\,\theta_L^E}\big)\Big)\\
&= \sum_{l=1}^{L} \partial_E \theta_l^E.
\end{aligned}
$$

Nun folgt die Behauptung durch Integration über E (wie in (6.7)). Damit ist gezeigt, dass $\langle\,\omega\,|\,\Gamma\,\rangle$ gleich der Summe der Windungszahlen der Eigenphasen ist. Nach (6.5) folgt damit sofort, dass (6.7) dem Bott-Maslov-Index entspricht. □

Der folgende Satz ist wichtig, um das Integral aus (6.7) zu berechnen:

Satz 6.3 *Sei die Abbildung* $E \mapsto \Phi^E$ *differenzierbar und* $\big[\Phi^E\big]_\sim \in \mathbb{L}_L^{\mathbb{C}}$. *Dann gilt:*

$$\Im\,\Bigg(\partial_E\ \log\Big(\det\big(\Pi([\Phi^E]_\sim)\big)\Big)\Bigg) = 2\ \mathrm{Spur}\ \Big(\big((\Phi^E)^*\Phi^E\big)^{-1}(\Phi^E)^*\mathcal{J}\ \big(\partial_E\Phi^E\big)\Big). \tag{6.8}$$

Beweis (vgl. [SB06]): Wie im Beweis von Satz 6.2 beginnt man mit der Identität

$$i\ \partial_E\ \log\Big(\det\big(\Pi([\Phi^E]_\sim)\big)\Big) = \mathrm{Spur}\Big(\big(\Pi([\Phi^E]_\sim)\big)^*\big(\partial_E\,\Pi([\Phi^E]_\sim)\big)\Big). \tag{6.9}$$

Dabei ist zu beachten, dass folgend Gleichung gilt:

$$i\ \partial_E\ \log\Big(\det\big([\Phi^E]_\sim\big)\Big) = -\Im\ \Big(\partial_E\ \log\Big(\det\big([\Phi^E]_\sim\big)\Big)\Big).$$

Daher folgt die Behauptung, wenn gezeigt werden kann, dass gilt:

$$-i\,\partial_E\ \log\Big(\det\big(\Pi([\Phi^E]_\sim)\big)\Big) = 2\ \mathrm{Spur}\ \Big(\big((\Phi^E)^*\Phi^E\big)^{-1}(\Phi^E)^*\mathcal{J}\ \big(\partial_E\Phi^E\big)\Big)$$

Jetzt führt man folgende $L \times L$ -Matrizen ein, die auch später von Bedeutung sein werden:

$$\phi_\pm^E = \begin{pmatrix}\mathbb{1} & \pm i\mathbb{1}\end{pmatrix}\Phi^E. \tag{6.10}$$

Wie immer schreibt man Φ^E als $\begin{pmatrix} a \\ b \end{pmatrix}$. Nach Definition von Π gilt damit

$$\Pi([\Phi^E]_\sim) = (a - ib)(a + ib)^{-1} = \phi_-^E(\phi_+^E)^{-1}.$$

Außerdem ist

$$
\begin{aligned}
(\phi_+^E)^*\phi_+^E &= (a+ib)^*(a+ib) \\
&= (a^*-ib^*)(a+ib) \\
&= a^*a + \underbrace{ia^*b - ib^*a}_{=0,\ \text{da } a^*b=b^*a} + b^*b \\
&= a^*a+b^*b \\
&\overset{\text{analog}}{=} (\phi_-^E)^*\phi_-^E \quad = \quad (\Phi^E)^*\Phi^E.
\end{aligned}
\tag{6.11}
$$

Ferner rechnet man nach, dass gilt:

$$
\left(\phi_\pm^E\right)^* \partial_E \phi_\pm^E = \left(\Phi^E\right)^* \partial_E \Phi^E \mp i \left(\Phi^E\right)^* \mathcal{J} \partial_E \Phi^E. \tag{6.12}
$$

Denn:

$$
\begin{aligned}
\left(\phi_\pm^E\right)^* \partial_E \phi_\pm^E &= \left(\begin{pmatrix} \mathbb{1} & \pm i\mathbb{1} \end{pmatrix} \Phi^E\right)^* \partial_E \begin{pmatrix} \mathbb{1} & \pm i\mathbb{1} \end{pmatrix} \Phi^E \\
&= (\Phi^E)^* \begin{pmatrix} \mathbb{1} \\ \mp i\mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbb{1} & \pm i\mathbb{1} \end{pmatrix} \partial_E \Phi^E \\
&= (\Phi^E)^* \begin{pmatrix} \mathbb{1} & \pm i\mathbb{1} \\ \mp i\mathbb{1} & \mathbb{1} \end{pmatrix} \partial_E \Phi^E \\
&= (\Phi^E)^* \left(\mathbb{1} \mp i\,\mathcal{J}\right) \partial_E \Phi^E \\
&= (\Phi^E)^* \partial_E \Phi^E \mp i\,(\Phi^E)^* \mathcal{J}\, \partial_E \Phi^E.
\end{aligned}
$$

Nun ist

$$
\begin{aligned}
\partial_E \Pi\left([\Phi^E]_\sim\right) &= \partial_E\left(\phi_-^E(\phi_+^E)^{-1}\right) \\
&= (\partial_E\phi_-^E)(\phi_+^E)^{-1} - \phi_-^E(\phi_+^E)^{-1}(\partial_E\phi_+^E)(\phi_+^E)^{-1}.
\end{aligned}
$$

Mit Hilfe der bisherigen Resultate kann man die gesuchte Spur auch anders ausdrücken, denn es gilt:

$$
\begin{aligned}
& \operatorname{Spur}\left(\left(\Pi([\Phi^E]_\sim)\right)^*\left(\partial_E\, \Pi([\Phi^E]_\sim)\right)\right) \\
= \quad & \operatorname{Spur}\left(\left(\phi_-^E(\phi_+^E)^{-1}\right)^*\left((\partial_E\phi_-^E)(\phi_+^E)^{-1} - \phi_-^E(\phi_+^E)^{-1}(\partial_E\phi_+^E)(\phi_+^E)^{-1}\right)\right) \\
= \quad & \operatorname{Spur}\Big(\underbrace{((\phi_+^E)^{-1})^*(\phi_-^E)^*\,(\partial_E\,\phi_-^E)(\phi_+^E)^{-1}}_{:=C_1} \\
& \qquad \underbrace{-\,\left(((\phi_+^E)^{-1})^*(\phi_-^E)^*\,\phi_-^E(\phi_+^E)^{-1}(\partial_E\phi_+^E)(\phi_+^E)^{-1}\right)}_{:=C_2}\Big).
\end{aligned}
$$

Um die gesuchte Spur zu berechnen, zerlegt man – wie oben schon angedeutet – diese Formel in zwei Teile (C_1 und C_2). Für den ersten Teil der gesuchten Spur ergibt sich aus der Zyklizität der Spur, dass

$$\text{Spur }(C_1) = \text{Spur}\Big(\big((\phi_+^E)^*(\phi_+^E)\big)^{-1}(\phi_-^E)^*\partial_E\phi_-^E\Big).$$

und für den zweiten Teil, C_2, erhält man:

$$\begin{aligned}\text{Spur }(C_2) &= -\text{ Spur }\Big(\underbrace{\big(\phi_+^E(\phi_+^E)^*\big)^{-1}\big((\phi_-^E)^*\ \phi_-^E\big)}_{=\mathbb{1}}(\phi_+^E)^{-1}(\partial_E\phi_+^E)\big)\\ &= -\text{ Spur}\Big((\phi_+^E)^{-1}\partial_E\phi_+^E\Big)\\ &= -\text{ Spur}\Big(\big((\phi_+^E)^*\ \phi_+^E\big)^{-1}(\phi_+^E)^*\ \partial_E\ \phi_+^E\Big).\end{aligned}$$

Nun setzt man die beiden Teile wieder zusammen und erhält:

$$\begin{aligned}&\text{Spur }(C_1+C_2)\\ =\ &\text{Spur }\Big(\big((\phi_+^E)^*\phi_+^E\big)^{-1}(\phi_-^E)^*\partial_E\phi_-^E - \big((\phi_+^E)^*\phi_+^E\big)^{-1}(\phi_+^E)^*\partial_E\phi_+^E\Big)\\ =\ &\text{Spur }\Big(\big((\Phi^E)^*\Phi^E\big)^{-1}\big((\phi_-^E)^*\partial_E\ \phi_-^E - (\phi_+^E)^*\partial_E\phi_+^E\big)\Big)\\ =\ &\text{Spur }\Big(\big((\Phi^E)^*\Phi^E\big)^{-1}\big((\Phi^E)^*\ \partial_E\ \Phi^E + i(\Phi^E)^*\ \mathcal{J}\partial_E\ \Phi^E\\ &\qquad -(\Phi^E)^*\ \partial_E\ \Phi^E + i(\Phi^E)^*\ \mathcal{J}\partial_E\ \Phi^E\big)\Big)\\ =\ &2\,i\text{ Spur }\Big(\big((\Phi^E)^*\Phi^E\big)^{-1}(\Phi^E)^*\ \mathcal{J}\partial_E\ \Phi^E\Big).\end{aligned}$$

Damit folgt die Behauptung. □

In diesem Kapitel wurde eine Formel für den Schnittindex eines geschlossenen Pfades Γ von Lagrange-Ebenen in $\mathbb{L}_L^{\mathbb{C}}$ und einer weiteren Lagrange-Ebene $[\Psi_\xi]_\sim$ hergeleitet, diese Formel ist wiederum mit den Windungszahlen verknüpft, wodurch man Gleichungen erhält, mit deren Hilfe von den Bott-Maslov-Index sehr einfach bestimmen kann.

Besonders wichtig für die Herleitung des Oszillationstheorems ist aber die Definition der singulären Zykel und die Charakterisierung der Schnitte mittels der Wronski-Matrix. Im nächsten Kapitel werden diese Ergebnisse auf die Sturm-Liouville-Probleme angewendet.

Kapitel 7

Lösungen des Sturm-Liouville-Problems

Mathematics is not only real, but it is the only reality.
That is that entire universe is made of matter, obviously.
And matter is made of particles. It's made of electrons and neutrons
and protons. So the entire universe is made out of particles.
Now what are the particles made of? They're not made out of anything.
The only thing you can say about the reality of an electron
is to cite the mathematical properties.
So there's a sense in which matter has completely dissolved
and what is left is just a mathematical structure.

Martin Gardner, [Ga94].

In diesem Kapitel wenden wir die Resultate aus den Kapiteln drei bis sechs auf die Sturm-Liouville-Probleme, die im zweiten Kapitel vorgestellt wurden, an. Dieses Kapitel wird ein Verfahren zur Bestimmung der Eigenwerte präsentieren. Der erste Teil dieses Kapitels ist an [SB06] angelehnt, im zweiten Abschnitt wird dann die Monotonie der Eigenphasen θ_l^E in E bewiesen, wobei man sich teilweise auf [Bo56] stützen kann. Eine genauere Untersuchung des Konvergenzverhaltens der Eigenphasen für $E \to -\infty$ folgt im achten Kapitel.

7.1 Dynamik der Lagrange-Ebenen

Bisher wurde schon gezeigt, dass die Lösungen der Gleichung $\mathcal{L}y = Ey$ den Lösungen der Gleichung $u_x' = \mathfrak{t}_x^E \, u_x$ entsprechen. Dabei ist

$$\mathfrak{t}_x^E = \begin{pmatrix} -p_x^{-1} q_x & p_x^{-1} \\ v_x - q_x^* p_x^{-1} q_x - E\mathbb{1} & q_x^* p_x^{-1} \end{pmatrix}.$$

Darüber hinaus wurde in Abschnitt 3.5 gezeigt, dass die Lösungen dieses Systems Lagrange-Ebenen $[\Phi_x^E]_\sim$ entsprechen. In Lemma 3.6 und Bemerkung 3.4 wurde erläutert, dass die Dynamik der Lösungen mit Hilfe von symplektischen Matrizen beschrieben werden kann, d. h. die Lösung an einem Punkt x, $x \in [0,\ 1]$, ist gegeben durch $[\Phi_x^E]_\sim = \mathfrak{T}_x^E \, [\Phi_0^E]_\sim$, wobei $[\Phi_0^E]_\sim \in \mathbb{L}_L^{\mathbb{C}}$ ist. Speziell gilt:

$$[\Phi_1^E]_\sim = \mathfrak{T}_1^E \, [\Phi_0^E]_\sim. \tag{7.1}$$

Mit Hilfe der Cayley- und der Möbius-Transformation haben wir das System

$$\Pi\big([\Phi_x^E]_\sim\big) = \Pi\big([\mathfrak{T}_x^E \Phi_0^E]_\sim\big) = \mathcal{C}\, \mathfrak{T}^E \, \mathcal{C}^* \cdot \Pi\left([\Phi_0^E]_\sim\right)$$

entwickelt, das die vorgegebenen Lösungspfade auf den unitären Raum überträgt. Mit Satz 5.4 haben wir die Wohldefiniertheit dieses Systems bewiesen, die Analytizität in E wurde in Satz 5.5 und Bemerkung 3.4 gezeigt.

Satz 7.1 *Es sei $(\mathcal{L}_\xi, \mathcal{D}(\mathcal{L}_\xi))$ der Sturm-Liouville-Operator zu den Randbedingungen $u_0 \in [\Psi_{\xi_0}]_\sim$, $u_1 \in [\Psi_{\xi_1}]_\sim$. Dabei seien $\Psi_{\xi_j} = \begin{pmatrix} \xi_j \\ \mathbb{1} \end{pmatrix}$ mit $\xi_j \in \mathrm{Her}(L,\ \mathbb{C})$ Repräsentanten von $[\Psi_{\xi_1}]_\sim$ ($j = 1,\ 2$). Dann gilt:*

$$\begin{aligned} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\xi \\ = \;& L - \mathrm{Rang}\Big(\big([\Psi_{\xi_1}]_\sim\big)^* \, \mathcal{J}\, [\Phi_1^E]_\sim\Big), \end{aligned}$$

wobei $[\Phi_1^E]_\sim$ durch (7.1) gegeben ist.

Beweis: Um die Eigenwerte des Operators $(\mathcal{L}_\xi, \mathcal{D}(\mathcal{L}_\xi))$ zu bestimmen, betrachtet man zunächst den Unterraum, der durch die Anfangsbedingung $[\Phi_0^E]_\sim = [\Psi_{\xi_0}]_\sim$ gegeben ist und folgt dann der Dynamik, die durch die Transfermatrix $\mathfrak{T}_x^E$ induziert wird, bis zum rechten Rand. Dann prüft man, ob der Unterraum $[\Phi_1^E]_\sim$, der sich auf diese Weise ergeben hat, einen nicht trivialen Schnitt mit dem Unterraum $[\Psi_{\xi_1}]_\sim$ (Randbedingung am rechten Rand) hat. Um die Schnitte zweier Lagrange-Ebenen zu charakterisieren, wurde in Definition 6.2 die Wronski-Matrix $\mathcal{W}$ zweier

Lagrange-Ebenen definiert. Lemma 6.1 hat gezeigt, dass $\dim\big(\Phi\, \mathbb{C}^L \,\cap\, \Psi_\xi\, \mathbb{C}^L\big) = \dim\Big(\ker\ \mathcal{W}\big(\Phi,\ \Psi_\xi\big)\Big)$ ist. Dieses Ergebnis nutzt man jetzt, um die Eigenwerte E von $(\mathcal{L}_\xi, \mathcal{D}(\mathcal{L}_\xi))$ mit ihrer Multiplizität zu bestimmen.

Da die Randbedingungen am linken Rand erfüllt sind (man wähle sie schlicht als Startbedingung!), ist $\mathrm{Rang}\Big(\big([\Psi_\xi]_\sim\big)^*\, \mathcal{J}\, [\Phi_0^E]_\sim\Big) = L$. Der Rang am rechten Rand ist dagegen durch $\mathrm{Rang}\,\Big(\big([\Psi_\xi]_\sim\big)^*\, \mathcal{J}[\Phi_1^E]_\sim\Big)$ gegeben. Insgesamt erhält man:

$$\begin{aligned} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\xi \\ = \quad & L - \mathrm{Rang}\Big(\big([\Psi_\xi]_\sim\big)^*\, \mathcal{J}\, [\Phi_1^E]_\sim\Big). \end{aligned} \tag{7.2}$$

Dies entspricht der Behauptung. □

Bemerkung 7.1 *Wenn man Probleme der Perturbationstheorie lösen möchte, kann es sinnvoll sein, das System unter einer symplektischen Basistransformation zu betrachten.*[1]

Im folgenden Satz zeigen wir Hauptresultat 4:

Satz 7.2 *Sei* $\varphi \in (0,\ 2\pi)$ *und* $\xi = -\cot\left(\frac{\varphi}{2}\right)\mathbb{1}$.*Dann gilt:*

$$\begin{aligned} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\varphi(1) \\ = \quad & \text{Multiplizität von } e^{i\varphi} \text{ als Eigenwert von } U_1^E. \end{aligned}$$

Beweis (vgl. [SB06]): Aus $\varphi \in (0,\ 2\pi)$ und $\xi = -\cot\left(\frac{\varphi}{2}\right)\mathbb{1}$ folgt sofort:

$$[\Psi_\xi]_\sim = \left[\Psi_{-\cot\left(\frac{\varphi}{2}\right)\mathbb{1}}\right].$$

[1] Dazu gibt man sich eine symplektische Matrix $\mathfrak{M} \in \mathfrak{SP}(2L, \mathbb{C})$ vor. Die Dynamik der Lösungen ist durch die Gleichung $([\Phi_x^E]_\sim)' = \mathfrak{M}\, \mathfrak{t}_x^E\, \mathfrak{M}^{-1}([\Phi_x^E]_\sim)$ gegeben. Damit gilt auch hier wieder, dass man zur Beschreibung der Lösungen $[\Phi_x^E]_\sim$ einen Pfad symplektischer Matrizen $\mathfrak{T}_x^E$ hat, für den gilt: $[\Phi_x^E]_\sim = \mathfrak{T}_x^E\, [\Phi_0^E]_\sim = [\mathfrak{T}_x^E\, \Phi_0^E]_\sim$. Dabei ist $[\Phi_0^E]_\sim$ eine Lagrange-Ebene. Für die Eigenwerte ergibt sich somit unter symplektischer Basistransformation:

$$\begin{aligned} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\phi(1) \\ = \quad & L - \mathrm{Rang}\,\Big(\big(\mathcal{M}\, [\Psi_\xi]_\sim\big)^*\, \mathcal{J}\, [\Phi_1^E]_\sim\Big). \end{aligned}$$

Den zugehörigen Sturm-Liouville-Operator bezeichne ich mit $\mathcal{L}_\varphi$. Für $\varphi = \pi$ ergeben sich Dirichlet-Randbedingungen.

Nach Satz 6.1 und (6.2) gilt:

$$\begin{array}{rl} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\varphi(1) \\ = & \text{Multiplizität von } e^{i\varphi} \text{ als Eigenwert von } U_1^E. \qquad \square \end{array} \tag{7.3}$$

Explizit gilt für Dirichlet-Randbedingungen:

$$\begin{array}{rl} & \text{Multiplizität von } E \text{ als Eigenwert von } \mathcal{L}_\varphi(1) \\ = & \text{Multiplizität von } e^{i\pi} \text{ als Eigenwert von } U_1^E. \end{array}$$

7.2 Monotonie und Transversalität

Im Zentrum dieses Abschnitts steht Satz 7.3, der zeigt, dass die Kurve $\Gamma = \left([\Phi_1^E]_\sim\right)_{E \in \mathbb{R}}$ transversal zu dem singulären Zykel $\mathbb{L}_L^\varphi$ ist und dass Γ ihn immer von der negativen zur positiven Seite durchquert.

Satz 7.3 *Es sei $U_1^E = \Pi\left([\Phi_1^E]_\sim\right)$, wobei $[\Phi_1^E]_\sim$ die Lösung des modifizierten Sturm-Liouville-Problems ist nach Definition 8.1 ist. Für $E \in \mathbb{R}$ gilt:*

$$\frac{1}{i}\left(U_1^E\right)^* \partial_E U_1^E > 0.$$

Beweis: Der Beweis ist angelehnt an den Beweis, den Bott (in [Bo56]) genutzt hat, um die Monotonie zu zeigen. Allerdings gilt Botts Beweis nur für die Fundamentalmatrizen aus $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ liegen. In dieser Arbeit wird dagegen das Verhalten der unitären Matrizen U_1^E betrachtet, das sich auf das Verhalten der Lagrange-Ebenen $[\Phi_1^E]_\sim$ zurückführen lässt. Dazu musste der Beweis von mir erweitert und modifiziert werden, und er birgt in diesem Sinne neue Erkenntnisse.

Für den Beweis betrachte man die schon aus (6.10) bekannten Matrizen $\phi_\pm^E = \left(\mathbb{1} \quad \pm i\mathbb{1}\right)\Phi^E$. Damit sind $\phi_\pm^E$ invertierbare $L \times L$ -Matrizen und es gilt $U^E = \phi_-^E(\phi_+^E)^{-1} = \left((\phi_-^E)^{-1}\right)^*(\phi_+^E)^*$. Insgesamt ergibt sich:

$$(U^E)^*\partial_E U^E = \left((\phi_+^E)^{-1}\right)^*\left((\phi_-^E)^*\partial_E\phi_-^E - (\phi_+^E)^*\partial_E\phi_+^E\right)\left(\phi_+^E\right)^{-1}.$$

Das bedeutet, für den Beweis genügt es zu zeigen:

$$\frac{1}{i}\left((\phi_-^E)^*\partial_E\phi_-^E - (\phi_+^E)^*\partial_E\phi_+^E\right) = 2\left(\Phi^E\right)^* \mathcal{J}\partial_E\Phi^E > 0.$$

Dazu betrachte man die Abbildung $(E,\, x) \mapsto \Phi_x^E$, wobei $[\Phi_x^E]_\sim$ nach wie vor die Lagrange-Ebene der Lösungen des Sturm-Liouville-Problems $\mathcal{L}y = Ey$ darstellt. Seien E und $h \in \mathbb{R}$ fest gewählt. Dann gilt für den Ausdruck $(\Phi_x^E)^* \mathcal{J} \Phi_x^{E+h}$, aufgefasst als Funktion in x:

$$\begin{aligned}
\Big((\Phi_x^E)^* \mathcal{J} \Phi_x^{E+h}\Big)' &= \big((\Phi_x^E)^*\big)' \mathcal{J} \Phi_x^{E+h} + (\Phi_x^E)^* \mathcal{J} (\Phi_x^{E+h})' \\
&= (\Phi_x^E)^* (\mathfrak{t}_x^E)^* \mathcal{J} \Phi_x^{E+h} + (\Phi_x^E)^* \mathcal{J} \mathfrak{t}_x^{E+h} \Phi_x^{E+h} \\
&= (\Phi_x^E)^* \big((\mathfrak{t}_x^E)^* \mathcal{J} + \mathcal{J} \mathfrak{t}_x^{E+h}\big) \Phi_x^{E+h}
\end{aligned}$$

Da $\mathfrak{t}_x^E$ in der symplektischen Lie-Algebra liegt, ist $(\mathfrak{t}_x^E)^* \mathcal{J} + \mathcal{J} \mathfrak{t}_x^E = 0$ für alle $x \in I$. Daraus folgt:

$$(\mathfrak{t}_x^E)^* \mathcal{J} + \mathcal{J} \mathfrak{t}_x^{E+h} = (\mathfrak{t}_x^E)^* \mathcal{J} + \mathcal{J} \mathfrak{t}_x^E + \mathcal{J} \begin{pmatrix} 0 & 0 \\ -h\mathbb{1} & 0 \end{pmatrix} = \mathcal{J} \begin{pmatrix} 0 & 0 \\ -h\mathbb{1} & 0 \end{pmatrix}.$$

Insgesamt erhält man:

$$\Big((\Phi_x^E)^* \mathcal{J} \Phi_x^{E+h}\Big)' = h\Big((\Phi_x^E)^* \mathcal{J} Q \Phi_x^{E+h}\Big),$$

wobei $Q = \begin{pmatrix} 0 & 0 \\ -\mathbb{1}_L & 0 \end{pmatrix}$ ist.

Nach dem Fundamentalsatz der Differential- und Integralrechnung gilt daher:[2]

$$(\Phi_a^E)^* \mathcal{J} \Phi_a^{E+h} - (\Phi_0^E)^* \mathcal{J} \Phi_0^{E+h} = h \int_0^a \mathrm{dx}\; (\Phi_x^E)^*\, R_x\, \Phi_x^{E+h}, \tag{7.4}$$

mit der Bezeichnung $R = \mathcal{J} Q = \begin{pmatrix} \mathbb{1}_L & 0 \\ 0 & 0 \end{pmatrix}$.

Aus der Randbedingung am linken Rand kann man entnehmen, dass es einen Repräsentanten Φ_0^E von $[\Phi_0^E]_\sim$ gibt, für den $\Phi_0^E = \begin{pmatrix} \xi \\ \mathbb{1} \end{pmatrix}$ ist. Daher ist

$$(\Phi_0^E)^* \mathcal{J} \Phi_0^{E+h} = \begin{pmatrix} \xi^* & \mathbb{1} \end{pmatrix} \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \mathbb{1} \end{pmatrix} = 0$$

(dabei beachte man, dass $\xi = \xi^*$ ist). Gemäß ihrer Definition gilt für die Repräsentanten Φ einer Lagrange-Ebene $[\Phi]_\sim$ stets $\Phi^* \mathcal{J} \Phi = 0$, d. h. $\big((\Phi_a^E)^* \mathcal{J} \Phi_a^E\big) = 0$

[2] Dabei ist das Integral komponentenweise definiert, d. h. für $T_x \in \mathrm{Mat}(L \times L, \mathbb{C})$ ist $\int \mathrm{dx}\ T_x$ definiert durch $\int \mathrm{dx}\ (T_x)_{i,j}$, für alle $i,\, j = 1,\ \ldots,\ L$.

für alle $a \in I$. Daher kann man in (7.4) den Term $(\Phi_0^E)^* \mathcal{J} \Phi_0^{E+h}$ durch $(\Phi_a^E)^* \mathcal{J} \Phi_a^E$ ersetzen. Man erhält also:

$$(\Phi_a^E)^* \mathcal{J} \Phi_a^{E+h} - (\Phi_a^E)^* \mathcal{J} \Phi_a^E = h \int_0^a \mathrm{dx}\ (\Phi_x^E)^*\ R_x\ \Phi_x^{E+h}.$$

Diese Gleichung teilt man durch h und geht dann zum Grenzwert $h \to 0$ über:

$$\lim_{h \to 0} \left(\frac{(\Phi_a^E)^* \mathcal{J} \Phi_a^{E+h} - (\Phi_a^E)^* \mathcal{J} \Phi_a^E}{h} \right) = \int_0^a \mathrm{dx}\ (\Phi_x^E)^*\ R_x\ \Phi_x^E,$$

d. h.

$$(\Phi_a^E)^*\ \mathcal{J}\ \partial_E \left(\Phi_a^E\right) = \int_0^a \mathrm{dx} (\Phi_x^E)^* R_x \Phi_x^E.$$

Jetzt gilt es noch zu zeigen, dass die linke Seite positiv definit ist. Sei dazu $\mathbb{C}^L \ni \rho \neq 0$. Dann gibt es ein eindeutiges $y_x \in \mathbb{C}^L$, für das $u_x = \Phi_x^E\ \rho$ gilt, d. h. y_x löst $\mathcal{L} y_x = E y_x$. Wertet man die Form auf der rechten Seite aus, erhält man:

$$\int_0^a \mathrm{dx}\ (\Phi_x^E \rho)^*\ R_x\ \Phi_x^E \rho = \int_0^a \mathrm{dx}\ \langle y_x | y_x \rangle.$$

Da $a \neq 0$ ist, ergibt sich, dass der Ausdruck auf der rechten Seite größer als 0 ist. Dies beweist die Behauptung. □

Obige Aussage ist offensichtlich für alle $x \in I$, d. h. für alle $\Pi\left(\Phi_x^E\right) = U_x^E$, gültig. Von besonderer Bedeutung ist der vorige Satz am rechten Rand, da man besonders an $\Pi\left(\Phi_1^E\right) = U_1^E$ interessiert ist.

Warum ist der vorige Satz so zentral für die Aussage dieser Arbeit? Auf der Basis dieses Satzes kann man die Monotonie der Eigenphasen in E beweisen. Man betrachte dazu die Abbildung $\mathbb{R} \ni E \mapsto U_1^E$. Dann gilt für das Spektrum von U_1^E nach (6.2): $\sigma\left(U_1^E\right) = \left\{ e^{i\theta_1^E},\ \ldots,\ e^{i\theta_L^E} \right\}$.

Satz 7.4 *Alle Eigenphasen $e^{i\theta_l^E}$ drehen sich (abhängig von der Energie E) im positiven Sinne.*

Beweis: Der Beweis basiert auf Satz 7.3. Seien u_l die zu jedem $e^{i\theta_l^E}$ gehörenden Eigenvektoren, d. h. $U_1^E |u_l\rangle = e^{i\theta_l^E} |u_l\rangle$ für alle $l = 1,\ \ldots,\ L$. Zudem seien die u_l so normiert, dass stets $\langle u_l | u_l \rangle = 1$ gilt.

Weiter gilt für alle l:

$$
\begin{aligned}
0 &< \left\langle u_l \left| \frac{1}{i} \left(U_1^E\right)^* \; \partial_E \, U_1^E \right| u_l \right\rangle \\
&= \frac{1}{i} \left\langle u_l \left| \left(U_1^E\right)^* \partial_E \; \left(e^{i\theta_l^E}\right) \right| u_l \right\rangle \\
&= \frac{1}{i} \left(U_1^E |u_l\rangle\right)^* \; i \, \partial_E(\theta_l^E) e^{i\theta_l^E} |u_l\rangle \\
&= \left(e^{i\theta_l^E} |u_l\rangle\right)^* \; \partial_E(\theta_l^E) e^{i\theta_l^E} |u_l\rangle \\
&= \langle u_l | e^{-i\theta_l^E} \; \partial_E(\theta_l^E) e^{i\theta_l^E} |u_l\rangle \\
&= \partial_E \, (\theta_l^E) \langle u_l | u_l \rangle \\
&= \partial_E \, (\theta_l^E).
\end{aligned}
$$

Damit ist die Behauptung bewiesen. □

Aus den Sätzen 7.3 und 7.4 ergibt sich Hauptresultat 5. Außerdem erhalten wir die folgende Aussage über das Spektrum von $(\mathcal{L}, \mathcal{D})$:

Folgerung 7.1
Diskretes Spektrum: *Das Spektrum von $(\mathcal{L}, \mathcal{D})$ versehen mit fixierten Randbedingungen Φ_{ξ_0}, Φ_{ξ_1} ist diskret. Denn da die Eigenphasen in E streng monoton wachsend sind, sind die Stellen, an denen die Eigenphasen den Zykel schneiden, diskret.*

Eigenwerte höchstens L -fach entartet: *Weil das Spektrum von U_1^E die L Eigenphasen $e^{i\theta_1^E}$, $\ldots$, $e^{i\theta_L^E}$ umfasst, können auch nur höchstens L von ihnen den singulären Zykel zu einem festen Zeitpunkt x treffen. Also hat $\mathcal{L}$ höchstens L -fach entartete Eigenwerte.*

Nachdem wir in diesem Kapitel das Oszillationstheorem hergeleitet und die Hauptresultate 4 und 5 bewiesen haben, beschäftigt sich das nächste Kapitel mit dem Verhalten der Eigenphasen für $E \to -\infty$.

Kapitel 8

Verhalten der Lösungen für $E \to -\infty$

Das Unendliche hat wie keine andere Frage
von jeher so tief das Gemt des Menschen bewegt;
das Unendliche hat wie kaum eine andere Idee
auf den Verstand so anregend und füruchtbar gewirkt;
das Unendliche ist aber auch wie kein anderer Begriff
so der Aufklärung bedrftig.

David Hilbert, [Hil64].

Die Problematik des Konvergenzverhaltens für $E \to -\infty$ wurde schon im vorhergehende Kapitel angerissen. Nun steht in diesem Kapitel die Frage im Mittelpunkt, was passiert, wenn E gegen $-\infty$ konvergiert. In diesem Kapitel wird der Sturm-Liouville-Operator stets als nach unten beschränkt voraussgesetzt, d. h. es existiert ein kleinster Eigenwertes $E_0 \in \mathbb{R}$. Dies entspricht einem ersten Schnitt mit dem singulären Zykel. Man kann erwarten, dass sich die Eigenphasen für sehr kleine Energien E nur sehr langsam drehen. Da die Eigenphasen monoton in E wachsen, werden sie sich mit größer werdender Energie immer schneller drehen. Um das Verhalten der Grenzwerte für $E \to -\infty$ expliziter zu studieren, bedient man sich eines Homotopiearguments, das auf Satz 7.3 und Satz 7.4 basiert.

Die Inhalte dieses Kapitels, sind – soweit ich das beurteilen kann – komplett neu.

8.1 Verhalten der Eigenphasen für $E \to -\infty$

Schon in [Bo56] wird ausgenutzt, dass die Eigenwerte des Sturm-Liouville-Problems $\mathcal{L}y = Ey$ geordnet sind und dass unter bestimmten Randbedingungen ein kleinster Eigenwert existiert.[1] Eine ähnliche Strategie verfolgt auch diese Arbeit.

Definition 8.1
Es sei $\lambda \in [0,\ 1]$. Man definiert die folgende Homotopie:

$$\lambda \mapsto \mathfrak{t}_x^{E,\lambda},$$

wobei $\mathfrak{t}_x^{E,\lambda}$ die folgende Gestalt hat:

$$\mathfrak{t}_x^{E,\lambda} = \begin{pmatrix} 0 & p_x^{-1} + \lambda\left(p_m \mathbb{1} - p_x^{-1}\right) \\ v_x - E\mathbb{1} + \lambda\left(v_m \mathbb{1} - v_x\right) & 0 \end{pmatrix}.$$

Dabei sind p_m und v_m reelle Konstanten, die die folgenden Eigeschaften haben:

$$\begin{array}{rcll} v_m \mathbb{1} & < & v_x & \text{für alle } x \in I \\ p_m \mathbb{1} & > & p_x^{-1} & \text{für alle } x \in I. \end{array}$$

Solche Konstanten existieren, da v_x und p_x^{-1} auf dem kompakten Intervall I stetig und somit beschränkt sind.

Mit Hilfe dieser Homotopie entsteht eine Schar von Differentialgleichungen erster Ordnung, die die Form $u_x' = \mathfrak{t}_x^{E,\lambda} u_x$ haben. Für alle $\lambda \in I$ kann man $\mathfrak{t}_x^{E,\lambda}$ auf die Form von $\mathfrak{t}_x^{E}$ bringen, indem man setzt:

$$\widetilde{p}_x^{-1} = p_x^{-1} + \lambda\left(p_m \mathbb{1} - p_x^{-1}\right), \quad \widetilde{v}_x = v_x + \lambda\left(v_m \mathbb{1} - v_x\right), \quad q \equiv 0.$$

Man beachte dabei, dass $\widetilde{p}_x^{-1}$ und $\widetilde{v}_x$ stetig sind. Damit erhält man also eine Schar von Lösungen des Sturm-Liouville-Problems. Diese Lösungen sind nach wie vor Lagrange-Ebenen, die man nun – da sie offensichtlich vom Parameter λ abhängen – mit $\Phi_x^{E,\lambda}$ bezeichnet.

Außerdem sei in diesem Kapitel stets $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ nach unten beschränkt. Da $q_x = 0$ für alle $x \in I$ ist, sind die Voraussetzungen an q von Lemma 2.3 ($q = q^*$ und $q \in C^1([0,\ 1], \mathrm{Mat}(L \times L, \mathbb{C}))$) erfüllt. Versehen mit Dirichlet-Randbedingungen erhält man nach Lemma 2.3 einen nach unten beschränkten Sturm-Liouville-Operator.

[1] Man beachte dabei, dass wegen der Selbstadjungiertheit von $\mathcal{L}$ alle Eigenwerte auf der reellen Achse liegen, wo eine Ordnungsstruktur gegeben ist.

Nun bedient man sich eines ähnlichen im Argumentes wie beim Beweis der Monotonie der Eigenphasen in E, um zu zeigen, dass die Eigenphasen auch in λ monoton wachsend sind.

Satz 8.1 *Es sei $U_1^{E,\lambda} = \Pi\left([\Phi_1^{E,\lambda}]_\sim\right)$, wobei $[\Phi_1^{E,\lambda}]_\sim$ die Lösung des modifizierten Sturm-Liouville-Problems ist nach Definition 8.1 ist. Für $E \in \mathbb{R}$ gilt:*

$$\frac{1}{i}\left(U_1^{E,\lambda}\right)^* \partial_\lambda U_1^{E,\lambda} > 0.$$

Beweis: Auch für diesen Beweis betrachte die Matrizen $\phi_\pm^E = \begin{pmatrix}\mathbb{1} & \pm i\mathbb{1}\end{pmatrix}\Phi_x^{E,\lambda}$. Damit sind $\phi_\pm^{E,\lambda}$ invertierbare $L \times L$-Matrizen und es gilt $U^{E,\lambda} = \phi_-^{E,\lambda}(\phi_+^{E,\lambda})^{-1} = \left((\phi_-^{E,\lambda})^{-1}\right)^*(\phi_+^{E,\lambda})^*$. Insgesamt ergibt sich:

$$(U^{E,\lambda})^*\partial_\lambda U^{E,\lambda} = \left((\phi_+^{E,\lambda})^{-1}\right)^*\left((\phi_-^{E,\lambda})^*\partial_\lambda\phi_-^{E,\lambda} - (\phi_+^{E,\lambda})^*\partial_\lambda\phi_+^{E,\lambda}\right)\left(\phi_+^{E,\lambda}\right)^{-1}.$$

Das bedeutet, es genügt für den Beweis die folgende positive Definitheit zu zeigen:

$$\frac{1}{i}\left((\phi_-^E)^*\partial_\lambda\phi_-^{E,\lambda} - (\phi_+^{E,\lambda})^*\partial_\lambda\phi_+^{E,\lambda}\right) = 2\left(\Phi^{E,\lambda}\right)^*\mathcal{J}\partial_\lambda\Phi^{E,\lambda} > 0$$

Dazu betrachte man die Abbildung $(\lambda,\, x) \mapsto \Phi_x^{E,\lambda}$. Seien nun $\lambda \in [0,\,1]$ und $h \in \mathbb{R}$ fest gewählt. Dann gilt für den Ausdruck $(\Phi_x^{E,\lambda})^*\mathcal{J}\Phi_x^{E,\lambda+h}$, aufgefasst als Funktion in x:

$$\begin{aligned}\left((\Phi_x^{E,\lambda})^*\mathcal{J}\Phi_x^{E,\lambda+h}\right)' &= \left((\Phi_x^{E,\lambda})^*\right)'\mathcal{J}\Phi_x^{E,\lambda+h} + (\Phi_x^{E,\lambda})^*\mathcal{J}(\Phi_x^{E,\lambda+h})' \\ &= (\Phi_x^{E,\lambda})^*(\mathfrak{t}_x^{E,\lambda})^*\mathcal{J}\Phi_x^{E,\lambda+h} + (\Phi_x^{E,\lambda})^*\mathcal{J}\mathfrak{t}_x^{E,\lambda+h}\Phi_x^{E,\lambda+h} \\ &= (\Phi_x^{E,\lambda})^*\left((\mathfrak{t}_x^{E,\lambda})^*\mathcal{J} + \mathcal{J}\mathfrak{t}_x^{E,\lambda+h}\right)\Phi_x^{E,\lambda+h}\end{aligned}$$

Da $\mathfrak{t}_x^{E,\lambda}$ in $\mathfrak{sp}(2L,\mathbb{C}$ liegt, ist $(\mathfrak{t}_x^{E,\lambda})^*\mathcal{J} + \mathcal{J}\mathfrak{t}_x^{E,\lambda} = 0$. Daraus folgt:

$$\begin{aligned}(\mathfrak{t}_x^{E,\lambda})^*\mathcal{J} + \mathcal{J}\mathfrak{t}_x^{E,\lambda+h} &= (\mathfrak{t}_x^{E,\lambda})^*\mathcal{J} + \mathcal{J}\mathfrak{t}_x^{E,\lambda} + h\mathcal{J}\begin{pmatrix}0 & p_m\mathbb{1} - p_x^{-1} \\ v_m\mathbb{1} - v_x & 0\end{pmatrix} \\ &= h\mathcal{J}\partial_\lambda\mathfrak{t}_x^{E,\lambda}.\end{aligned}$$

Insgesamt erhält man:

$$\left((\Phi_x^{E,\lambda})^*\mathcal{J}\Phi_x^{E,\lambda+h}\right)' = (\Phi_x^{E,\lambda})^* h\mathcal{J}\partial_\lambda\mathfrak{t}_x^{E,\lambda}\Phi_x^{E,\lambda+h}.$$

Daher gilt:

$$(\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda+h} - (\Phi_0^{E,\lambda})^* \mathcal{J} \Phi_0^{E,\lambda+h} = h \int_0^a \mathrm{d}x (\Phi_x^{E,\lambda})^* \, \mathcal{J} \partial_\lambda \mathsf{t}^{E,\lambda} \, \Phi_x^{E,\lambda+h}. \tag{8.1}$$

Wegen der Anfangsbedingung spannen $\Phi_0^{E,\lambda}$ und $\Phi_0^{E,\lambda+h}$ die gleiche Lagrange-Ebene auf, und es gilt:

$$(\Phi_0^{E,\lambda})^* \mathcal{J} \Phi_0^{E,\lambda+h} = 0.$$

Da außerdem per definitionem $\left((\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda}\right) = 0$ für alle $a \in I$ ist, kann man in (8.1) den Term $(\Phi_0^{E,\lambda})^* \mathcal{J} \Phi_0^{E,\lambda+h}$ durch $\left((\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda}\right)$ ersetzen. Man erhält:

$$(\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda+h} - (\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda} = h \int_0^a \mathrm{d}x (\Phi_x^{E,\lambda})^* \, \mathcal{J} \partial_\lambda \mathsf{t}^{E,\lambda} \, \Phi_x^{E,\lambda+h}.$$

Diese Gleichung teilt man durch h und geht dann zum Grenzwert $h \to 0$ über:

$$\lim_{h \to 0} \left(\frac{(\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda+h} - (\Phi_a^{E,\lambda})^* \mathcal{J} \Phi_a^{E,\lambda}}{h} \right) = \int_0^a \mathrm{d}x (\Phi_x^{E,\lambda})^* \, \mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} \, \Phi_x^{E,\lambda},$$

bzw.

$$(\Phi_a^{E,\lambda})^* \, \mathcal{J} \, \partial_\lambda \left(\Phi_a^{E,\lambda}\right) = \int_0^a \mathrm{d}x (\Phi_x^{E,\lambda})^* \, \mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} \, \Phi_x^{E,\lambda}.$$

Dieser Ausdruck ist positiv definit. Um dies einzusehen, betrachte man die Matrix $\mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda}$. Es gilt:

$$\mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} = \begin{pmatrix} v_x - v_m \mathbb{1} & 0 \\ 0 & p_m \mathbb{1} - p_x^{-1} \end{pmatrix}.$$

Die Konstanten v_m und p_m waren genau so gewählt, dass gilt:

$$v_x - v_m \mathbb{1} > 0 \quad \text{und} \quad p_m \mathbb{1} - p_x^{-1} > 0 \qquad (x \in I).$$

Daher ist für alle $x \in I$

$$\mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} > 0.$$

Sei nun $\mathbb{C}^L \ni \rho \neq 0$. Dann gibt es ein eindeutiges $y_x \in \mathbb{C}^L$, für das $u_x = \Phi_x^{E,\lambda} \rho$ gilt, d. h. y_x löst $\mathcal{L} y_x = E y_x$. Wertet man diese Form aus, erhält man:

$$\int_0^a \mathrm{d}x \, (\Phi_x^{E,\lambda} \rho)^* \, \mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} \, \Phi_x^E \rho = \int_0^a \mathrm{d}x \, \langle y_x | \mathcal{J} \partial_\lambda \mathsf{t}_x^{E,\lambda} | y_x \rangle.$$

Da $a > 0$ ist, ergibt sich aus der positiven Definitheit von $\mathcal{J} \partial_\lambda \mathsf{t}^{E,\lambda}$, dass der Ausdruck auf der rechten Seite größer als 0 ist. Dies beweist die Behauptung. □

Bemerkung 8.1 *Ist $\sigma\left(U_1^E\right) = \left\{e^{i\theta_1^{E,\lambda}}, \ldots, e^{i\theta_L^{E,\lambda}}\right\}$, ergibt sich analog zum Beweis von Satz 7.4, dass gilt:*

$$\partial_\lambda \theta_l^E > 0,$$

d. h. alle Eigenphasen drehen sich mit wachsendem λ im positiven Sinne.

Um eine Aussage über das Verhalten der Eigenphasen für $\lambda = 0$ zu machen, beschäftigt man sich zunächst mit ihrem Verhalten für $\lambda = 1$, da man in diesem Falle ein System von Differentialgleichungen mit konstanten Koeffizienten hat.

Satz 8.2 (Lösungen von $u'_x = \mathfrak{t}_x^{E,1} u_x$)
Die Lösungen des Problems $u'_x = \mathfrak{t}_x^{E,1} u_x$ unter der Anfangsbedingung $u_0 \in [\Phi_0^{E,1}]_\sim$ sind gegeben durch:

$$e^{\mathfrak{t}_x^{E,1} x}[\Phi_0^{E,1}]_\sim = \frac{1}{2}\begin{pmatrix} \left(e^{\tau^E x} + e^{-\tau^E x}\right)\mathbb{1} & \frac{p_m}{\tau^E}\left(e^{\tau^E x} - e^{-\tau^E x}\right)\mathbb{1} \\ \frac{\tau^E}{p_m}\left(e^{\tau^E x} - e^{-\tau^E x}\right)\mathbb{1} & \left(e^{\tau^E x} + e^{-\tau^E x}\right)\mathbb{1} \end{pmatrix}[\Phi_0^{E,1}]_\sim,$$

wobei $\tau^E := \sqrt{p_m(v_m - E)}$ ist.

Beweis: Der Beweis basiert auf den Methoden zur Lösung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten, die in Anhang C vorgestellt werden. Die Lösungen des Problems $u'_x = \mathfrak{t}_x^{E,1} u_x$ zur Anfangsbedingung $u_0 \in [\Phi_0^{E,1}]_\sim$ sind nach Satz C.4 gegeben durch $e^{\mathfrak{t}_x^{E,1} x}\,[\Phi_0^{E,1}]_\sim$.

Da $\mathfrak{t}_x^{E,1} = \begin{pmatrix} 0 & p_m\mathbb{1} \\ (v_m - E)\mathbb{1} & 0 \end{pmatrix}$ ist, sind $\pm\sqrt{p_m(v_m - E)}$ die Eigenwerte von $\mathfrak{t}_x^{E,1}$. Eine Basis aus Eigenvektoren ist daher von der Gestalt

$$\mathcal{B} = \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \frac{\sqrt{p_m(v_m-E)}}{p_m}\mathbb{1} & -\frac{\sqrt{p_m(v_m-E)}}{p_m}\mathbb{1} \end{pmatrix}.$$

Somit folgt:

$$\mathcal{B}^{-1} = \frac{1}{2}\begin{pmatrix} \mathbb{1} & \frac{p_m}{\sqrt{p_m(v_m-E)}}\mathbb{1} \\ \mathbb{1} & -\frac{p_m}{\sqrt{p_m(v_m-E)}}\mathbb{1} \end{pmatrix}.$$

Nun kann man die Lösungen explizit berechnen:

$$
\begin{aligned}
& e^{\mathfrak{t}_x^{E,1}\,x}\,[\Phi_0^{E,1}]_\sim \\
= \; & \mathcal{B}\begin{pmatrix} e^{\sqrt{p_m(v_m-E)}x}\mathbb{1} & 0 \\ 0 & e^{-\sqrt{p_m(v_m-E)}x}\mathbb{1} \end{pmatrix}\mathcal{B}^{-1}[\Phi_0^{E,1}]_\sim \\
= \; & \frac{1}{2}\begin{pmatrix} \left(e^{\tau^E x}+e^{-\tau^E x}\right)\mathbb{1} & \frac{p_m}{\tau^E}\left(e^{\tau^E x}-e^{-\tau^E x}\right)\mathbb{1} \\ \frac{\tau^E}{p_m}\left(e^{\tau^E x}-e^{-\tau^E x}\right)\mathbb{1} & \left(e^{\tau^E x}+e^{-\tau^E x}\right)\mathbb{1} \end{pmatrix}[\Phi_0^{E,1}]_\sim,
\end{aligned}
$$

wobei $\tau^E := \sqrt{p_m(v_m - E)}$ ist. $\square$

Satz 8.3 (Verhalten von $U_1^{E,1}$ für $E \to -\infty$)
Es sei $[\Phi_1^{E,1}]_\sim$ die Lösung des Problems $u_x' = \mathfrak{t}_x^{E,1} u_x$ zur Anfangsbedingung $u_0 \in [\Phi_0^{E,1}]_\sim$ am rechten Rand ($x = 1$). Dann gilt:

$$
\lim_{E\to-\infty}\left(U_1^E\right) = \lim_{E\to-\infty}\left(\Pi\left([\Phi_1^E]_\sim\right)\right) = -\mathbb{1}.
$$

Beweis: Sei $\begin{pmatrix} a \\ b \end{pmatrix}$ ein Repräsentant von $[\Phi_0^{E,1}]_\sim$. Nach Satz 8.2 ist ein Repräsentant der Lösung am rechten Rand gegeben durch

$$
\Phi_1^{E,1} = \frac{1}{2}\begin{pmatrix} e^{\tau^E}\left(a + \frac{p_m}{\tau^E}b\right) + e^{-\tau^E}\left(a - \frac{p_m}{\tau^E}b\right) \\ e^{\tau^E}\left(\frac{\tau^E}{p_m}a + b\right) + e^{-\tau^E x}\left(-\frac{\tau^E}{p_m}a + b\right) \end{pmatrix},
$$

wobei $\tau^E := \sqrt{p_m(v_m - E)}$ gilt. Da $|v_m| < \infty$ ist, gibt es ein $E_0 \in (-\infty, 0)$ so, dass $v_m - E$ für alle $E < E_0$ positiv ist. Daher kann man ohne Einschränkung annehmen, dass die Wurzel reell ist.

Mit Definition 4.3 erhält man:

$$
\begin{aligned}
& \Pi\left(\Phi_1^E\right) \\
= & \left(e^{\tau^E}\left(a + \frac{p_m}{\tau^E}b - i\left(\frac{\tau^E}{p_m}a + b\right)\right)\right)\left(e^{\tau^E}\left(a + \frac{p_m}{\tau^E}b + i\left(\frac{\tau^E}{p_m}a + b\right)\right)\right)^{-1} \\
& + \mathcal{O}(e^{-\tau^E}) \\
= & \left(a + \frac{p_m}{\tau^E}b - i\left(\frac{\tau^E}{p_m}a + b\right)\right)\left(a + \frac{p_m}{\tau^E}b + i\left(\frac{\tau^E}{p_m}a + b\right)\right)^{-1} + \mathcal{O}(e^{-\tau^E})
\end{aligned}
$$

Da $\tau^E := \sqrt{p_m(v_m - E)}$ ist, und da p_m und v_m konstante reelle Zahlen sind, gilt:

$$
\Pi\left(\Phi_1^E\right) = \left(-i\left(\frac{\sqrt{p_m(v_m - E)}}{p_m}a\right)\right)\left(i\left(\frac{\sqrt{p_m(v_m - E)}}{p_m}a\right)\right)^{-1} + \mathcal{O}(1).
$$

Insgesamt ergibt sich:

$$U_1^{E,1} = \Pi\left(\Phi_1^{E,1}\right) \to -\mathbb{1} \qquad \text{für } E \to -\infty. \qquad \square$$

Bemerkung 8.2 *Der Operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ sei nach unten beschränkt ist und $\lambda \in [0,\ 1]$ fest. Dann existiert ein $E_0 \in \mathbb{R}$, für das gilt: Für alle $E \in (-\infty,\ E_0)$ ist nur die triviale Lösung eine Lösung der Gleichung $\mathcal{L}y = Ey$.*

Wählt man etwa Dirichlet-Randbedingungen, *so tritt ein Eigenwert E genau dann auf, wenn eine Eigenphase $\theta_l^{E,\lambda} = \pi \bmod (2\pi)$ ist.*[2] *In Intervallen $[E_1,\ E_2]$ ohne Eigenwert bewegen sich die Phasen $\theta_l^{E,\lambda}$ nur im Bereich $(-\pi,\ \pi) \bmod (2\pi)$.*

Bemerkung 8.3 *Es sei E_0 der kleinste Eigenwert eines nach unten beschränkten Sturm-Liouville-Operators. Nach Bemerkung 8.2 kann man die Eigenphasen $\theta_l^{E,\lambda}$ so wählen, dass für $E \in (-\infty,\ E_0)$ und festes $\lambda \in [0,\ 1]$ gilt:*

$$\theta_l^{E,\lambda} \in (-\pi,\ \pi) \qquad \text{für alle } l = 1,\ \ldots,\ L.$$

Da nach Satz 8.3 $\lim_{E\to-\infty} U_1^{E,1} = -\mathbb{1}$ ist, gilt aufgrund der Monotonie der Phasen in E in diesem Falle: $\theta_l^{E,1} \to -\pi$ für $E \to -\infty$.

Satz 8.4 *Es sei $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ nach unten beschränkt, $\lambda \in [0, 1]$ fest und $\sigma\left(U_1^{E,\lambda}\right) = \left\{e^{i\theta_1^{E,\lambda}},\ \ldots,\ e^{i\theta_L^{E,\lambda}}\right\}$. Wählt man jedes $\theta_l^{E,\lambda}$ derart, dass $\theta_l^{E,\lambda}$ in $(-\pi,\ \pi)$ ist, so existiert der Grenzwert*

$$\lim_{E\to-\infty} \theta_l^{E,\lambda} \qquad \text{für alle } l = 1,\ \ldots,\ L.$$

Beweis: Zunächst betrachtet man die Abbildung $e_l\ :\ (-\infty, E_0) \ni E \mapsto \theta_l^{E,\lambda}$. Nach Satz 7.4 drehen sich die Eigenphasen im positiven Sinne, d. h. alle Abbildungen e_l sind monoton wachsend. Außerdem ist nach Folgerung 8.2 alle $\theta_l^{E,\lambda}$ auf den Bereich $(-\pi,\ \pi) \bmod (2\pi)$ beschränkt. Wählt man nun jedes $\theta_l^{E,\lambda}$ $(l = 1,\ \ldots,\ L)$ so, dass $\theta_l^{E,\lambda}$ nur in $(-\pi,\ \pi)$ sein darf, ist $e_l(E)$ darüber hinaus beschränkt. Das bedeutet, dass der Grenzwert $\lim_{E\to-\infty} \theta_l^{E,\lambda}$ existiert. $\square$

Nun kann man zeigen, dass auch der Grenzwert der unitären Matrix des Problems $u'_x = \mathfrak{t}_x^{E,0} u_x$ existiert. Dessen Lösungen entsprechen Lösungen des Sturm-Liouville-Problems $(p_x y'_x)' + v_x y_x = E y_x$.

[2]Dies entspricht genau dem Punkt -1 auf dem Rand des Einheitskreises.

Satz 8.5 (Verhalten von $U_1^{E,0}$)
Es sei $[\Phi_1^{E,0}]_\sim$ *die Lösung des Problems* $u'_x = \mathfrak{t}^{E,0} u_x$ *zur Anfangsbedingung* $u_0 \in [\Phi_0^{E,0}]_\sim$*. Dann ist*

$$\lim_{E\to-\infty} U_1^{E,0} = \lim_{E\to-\infty} \Pi\left(\left[\Phi_1^{E,0}\right]_\sim\right) = -\mathbb{1}.$$

Beweis: Es sei $\theta_l^{E,\lambda} \in (-\pi,\ \pi)$ für alle $\lambda \in [0,\ 1]$. Der Operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ hat nach Voraussetzung für alle λ eine untere Schranke. D. h. für alle $x \in I$ und $\lambda \in [0,\ 1]$ fest gilt:

$$\text{Es existiert ein } E_0^\lambda \in (-\infty,\ 0), \text{ so dass } \mathcal{L} \geq E_0^\lambda \text{ ist.}$$

Es sei nun $E_0 = \inf_{\lambda \in [0,\,1]} \left(E_0^\lambda\right)$. Da für alle $\lambda \in I$ gilt: $|E_0^\lambda| < \infty$, ist auch $|E_0| < \infty$. Somit folgt für alle $E < E_0$ und für alle $\lambda \in [0,\ 1]$, dass es keinen Schnitt mit dem singulären Zykel gibt. Für die Eigenphasen $\theta_l^{E,\lambda}$ ergibt sich also:

$$\theta_l^{E,\lambda} \in (-\pi,\pi) \mod (2\pi) \qquad \text{für alle } \lambda \in [0,\ 1] \text{ und } E < E_0. \tag{8.2}$$

Nun macht man von der Wahlfreiheit gebrauch und wählt alle Eigenphasen so, dass für $E < E_0$ und $\lambda \in [0,\ 1]$ gilt: $\theta_l^{E,\lambda} \in (-\pi,\ \pi)$. Nach Bemerkung 8.3 gilt (für $\lambda = 1$): $\lim_{E\to-\infty} \theta_l^{E,1} = -\pi$. Das bedeutet, dass für alle $\epsilon > 0$ ein E_1 existiert, so dass für alle $E < E_1$

$$\theta_l^{E,1} \in (-\pi,\ -\pi+\epsilon) \text{ ist} \qquad (l = 1,\ \ldots,\ L).$$

Nach Bemerkung 8.1 sind die Eigenphasen aber monoton wachsend in λ (d. h. $\partial_\lambda \theta_l^{E,0} > 0$), daher gilt auch

$$\lim_{E\to-\infty} \theta_l^{E,0} \in (-\pi,\ -\pi+\epsilon) \qquad (l = 1,\ \ldots,\ L). \tag{8.3}$$

Dabei beachte man, dass die Eigenphasen in 8.2 genau so bestimmt wurden, dass für $E < E_0$ und für alle $\lambda \in [0,\ 1]$ gilt: $\theta_l^{E,\lambda} \in (-\pi,\pi)$. Da (8.3) für alle $\epsilon > 0$ gilt, ergibt sich:

$$\lim_{E\to-\infty} \theta_l^{E,0} = \lim_{E\to-\infty} \theta_l^{E} = -\pi \qquad (l = 1,\ \ldots,\ L).$$

Daraus folgt die Behauptung. □

Dieser Satz hat das sechste Hauptresultat gezeigt und ist gleichzeitig der Abschluss des Hauptteils dieser Diplomarbeit.

Anhang A

Hermitesch symplektische Geometrie

Während im Hauptteil der Arbeit nur die symplektische Standardform betrachtet wird, erläutert dieser Anhang erläutert die symplektische Geometrie auf allgemeinen komplexen Vektorräumen. Diese Details mögen eher technisch scheinen, sie sind aber für das Verständnis der hermitesch symplektischen Geometrie wesentlich. Dieser Anhang ist in weiten Teilen eine Zusammenfassung der Ergebnisse, die es zur symplektischen Geometrie gibt. Dabei habe ich mich vor allem auf [deG97], [deG06], [Be98] sowie auf [Ar67], [Ar85] und [Ar00] gestützt.

A.1 Der hermitesch symplektische Raum

Definition A.1 *Ein Paar* $(\mathcal{V},\ \omega)$ *heißt* hermitesch symplektischer Raum, *wenn* $\mathcal{V}$ *ein komplexer Vektorraum und* ω *eine schiefsymmetrische, nichtdegenerierte Bilinearform auf* $\mathcal{V}$ *ist. Man bezeichnet* ω *dann als* hermitesch symplektische Form.

D. h. für $\nu,\ \nu_1,\ \nu_2 \in \mathcal{V}$ *gilt:*

- Schiefsymmetrie:

$$\omega(\nu_1,\nu_2) = -\omega(\nu_2,\nu_1), \qquad \text{und daher} \qquad \omega(\nu,\nu) = 0 \text{ für alle } \nu \in \mathcal{V}$$

- Nichtdegeneriertheit:

$$\omega(\nu_1,\nu_2) = 0 \text{ für alle } \nu_1 \in \mathcal{V} \qquad \Rightarrow \qquad \nu_2 = 0.$$

Im hermitesch symplektischen Raum definiert man die *hermitesch symplektische Orthogonalität* analog zur üblichen Definition für Skalarprodukte wie folgt:

Definition A.2 *Seien ν_1 und $\nu_2 \ \in \mathcal{V}$. Diese beiden Vektoren sind* hermitesch symplektisch orthogonal *zueinander, wenn gilt:*

$$\omega(\nu_1, \nu_2) = 0.$$

Die Relation, die durch diese Orthogonalität definiert ist, ist reflexiv (da wegen der Schiefsymmetrie $\omega(\nu, \nu) = 0$ gilt) und symmetrisch.

Lemma A.1 *Ein hermitesch symplektischer Raum ist immer von gerader Dimension.*

Beweis (vgl. [Be98, Kap. 1, S. 9/10]): In dem hier vorgestellten Beweis wird eine etwas allgemeinere Aussage bewiesen, die aber später noch von Nutzen sein wird: Falls ω eine hermitesch symplektische Form auf einem komplexen Vektorraum $\mathcal{V}$ ist und Rang $(\omega) = m$ ist, so gibt es ein $L \ \in \mathbb{N}$ für das gilt: $m = 2L$. Außerdem gibt es eine Basis $\{e_1, \ \ldots, \ e_{2L}\}$ von $\mathcal{V}$, bezüglich derer gilt:

$$\omega_e = \begin{pmatrix} 0 & -\mathbb{1}_L \\ \mathbb{1}_L & 0 \end{pmatrix}.$$

Da $\omega \neq 0$ ist, muss es zwei Vektoren e_1 und e_{L+1} in $\mathcal{V}$ geben, für die $\omega(e_1, e_{L+1}) \neq 0$ gilt. Indem man gegebenenfalls e_1 mit einem skalaren Faktor mulitpliziert, kann man o. B. d. A. annehmen, dass $\omega(e_1, e_{L+1}) = -1$ ist. Da ω als hermitesch symplektische Form schiefsymmetrisch ist, gilt darüber hinaus: $\omega(e_1, e_1) = \omega(e_{L+1}, e_{L+1}) = 0$ und $\omega(e_{L+1}, e_1) = -\omega(e_1, e_{L+1}) = 1$. Die Matrix $\widetilde{\omega}$, die nur die von e_1 und e_{L+1} aufgespannte Ebene $\mathcal{V}_1$ beschreibt, ist also durch $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ gegeben.

Sei nun $\mathcal{V}_1^\omega$ das hermitesch symplektisch orthogonale Komplement zu $\mathcal{V}_1$, d. h.

$$\mathcal{V}_1^\omega = \{\nu \ \in \ \mathcal{V} : \ \omega(\nu, \mu) = 0 \text{ für alle } \mu \ \in \ \mathcal{V}_1\}.$$

Offensichtlich ist $\mathcal{V}_1^\omega \cap \mathcal{V}_1 = \{0\}$ und $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_1^\omega$, denn für $\nu \ \in \ \mathcal{V}$ gilt

$$\nu - \omega(\nu, e_{L+1})e_1 + \omega(\nu, e_1)e_{L+1} \ \in \ \mathcal{V}_1^\omega.$$

Denn mit zwei komplexen Zahlen α und β gilt:

$$
\begin{aligned}
& \omega\Big(\nu-\omega(\nu,e_{L+1})e_1+\omega(\nu,e_1)e_{L+1},\ \alpha\, e_1+\beta\, e_{L+1}\Big)\\
= \ & \alpha\,\omega\Big(\nu-\omega(\nu,e_{L+1})e_1+\omega(\nu,e_1)e_{L+1},\ e_1\Big)\\
& +\beta\,\omega\Big(\nu-\omega(\nu,e_{L+1})e_1+\omega(\nu,e_1)e_{L+1},\ e_{L+1}\Big)\\
= \ & \alpha\Big(\omega(\nu,e_1)+\omega(\nu,e_{L+1})\underbrace{\omega(e_1,e_1)}_{=0}-\omega(\nu,e_1)\underbrace{\omega(e_{L+1},e_1)}_{=1}\Big)\\
& +\beta\Big(\omega(\nu,e_{L+1})+\omega(\nu,e_{L+1})\underbrace{\omega(e_1,e_{L+1})}_{=-1}-\omega(\nu,e_1)\underbrace{\omega(e_{L+1},e_{L+1})}_{=0}\Big)\\
= \ & 0.
\end{aligned}
$$

Jetzt wird das Verfahren für $\omega \neq 0$ auf $\mathcal{V}_1^{\omega}$ wiederholt, indem man e_2 und e_{L+2} wählt, die $\omega(e_2, e_{L+2}) = -1$ erfüllen.

Induktiv folgt somit die Behauptung und ω_e hat die gewünschte Form. $\square$

A.2 Lagrange-Ebenen

Definition A.3 (Lagrange-Ebenen)
Seien nun $\widetilde{\mathcal{V}}$, $\mathcal{V}_{coiso}$, $\mathcal{V}_{iso}$ und Φ Unterräume des hermitesch symplektischen Vektorraums $\mathcal{V}$. Dann definiert man:

$$\widetilde{\mathcal{V}}^{\,\omega} = \{\nu_1 \in \ \mathcal{V} \ : \omega(\nu_1,\nu_2) = 0 \text{ für alle } \nu_2 \ \in \ \widetilde{\mathcal{V}}\}.$$

Mit Hilfe dieser Definition kann man Unterräume von $\mathcal{V}$ wie folgt klassifizieren:

- *Ein Unterraum $\mathcal{V}_{iso} \subset \mathcal{V}$ heißt* isotrop, *falls $\mathcal{V}_{iso} \subset \mathcal{V}_{iso}^{\,\omega}$ ist.*
- *Ein Unterraum $\mathcal{V}_{coiso} \subset \mathcal{V}$ heißt* coisotrop, *falls $\mathcal{V}_{coiso}^{\,\omega} \subset \mathcal{V}_{coiso}$ ist.*
- *Falls ein Unterraum $\Phi \subset \mathcal{V}$ sowohl isotrop als auch coisotrop ist (d. h. $\Phi^{\,\omega} = \Phi$), nennt man Φ eine* ***Lagrange-Ebene****.*

Wie schon in obiger Definition angedeutet, werden Lagrange-Ebenen in dieser Arbeit immer mit Φ bezeichnet.

Aus der Nichtdegeneriertheit der hermitesch symplektischen Form ω folgt darüber hinaus sofort, dass in einem hermitesch symplektischen Vektorraum $\mathcal{V}$ der einzige Vektor, der zu allen anderen Vektoren orthogonal ist, der 0-Vektor ist.

Lemma A.2 *Sei $\mathcal{U}$ ein linearer Unterraum des hermitesch symplektischen, endlich dimensionalen Vektorraums $\mathcal{V}$, der mit der hermitesch symplektischen Form ω ausgestattet ist. Dann gilt für $\mathcal{U}^{\,\omega}$:*

$$\dim\ \mathcal{U} + \dim\ \mathcal{U}^{\,\omega} = \dim\ \mathcal{V}. \tag{A.1}$$

Beweis (vgl. [deG06, Kap. 1, S. 11/12]): Die Linearität des Unterraums $\mathcal{U}^{\,\omega}$ ist sofort klar.
Um die Dimensionsformel zu beweisen, definiert man eine lineare Form $\iota\colon \mathcal{V} \to \mathcal{U}$, die jedem $\nu_{\mathcal{U}} \in \mathcal{U}$ die lineare Form

$$\iota_v(\nu_{\mathcal{U}}) = \omega(v, \nu_{\mathcal{U}})$$

zuordnet. Da ω nichtdegeneriert und $\dim \mathcal{V} < \infty$ ist, definiert ι einen Isomorphismus. Sei nun $\{e_1, \dots, e_k\}$ eine Basis von $\mathcal{U}$. Dann ist

$$\mathcal{U}^{\,\omega} = \bigcap_{j=1}^{k} \ker\big(\iota(e_j)\big).$$

Daher ist $\mathcal{U}^{\,\omega}$ durch k linear unabhängige Gleichungen definiert, und somit gilt:

$$\dim(\mathcal{U}^{\,\omega}) = \dim(\mathcal{V}) - k = \dim(\mathcal{V}) - \dim(\mathcal{U}).$$

Daraus folgt sofort die Behauptung. □

Damit erhält man:

Lemma A.3 *Sei $\mathcal{V}$ ein Vektorraum,* $\dim(\mathcal{V}) = 2L,\ L \in \mathbb{N}$. *Ein Unterraum Φ von $\mathcal{V}$ ist genau dann eine Lagrange-Ebene, wenn gilt:*

$$\dim(\Phi) = L \qquad \text{und} \qquad \omega \text{ verschwindet identisch auf } \Phi.$$

Der Begriff der Lagrange-Ebene wird später von großer Bedeutung sein und im Laufe dieser Arbeit noch näher erläutert werden, doch zunächst:

A.3 Die Symplektische Gruppe

Definition A.4 *Ein* hermitesch symplektischer Automorphismus *auf* $(\mathcal{V}, \omega)$ *ist eine lineare Abbildung $s : \mathcal{V} \to \mathcal{V}$, so dass $s^*\,\omega = \omega$ ist, d. h.:*

$$\omega\big(s(\nu_1), s(\nu_2)\big) = \omega(\nu_1, \nu_2)\ \forall\ \nu_1,\ \nu_2\ \in\ \mathcal{V}.$$

Mit dieser Definition gilt:

Satz A.1 *Die hermitesch symplektischen Automorphismen* $(\mathcal{V},\omega) \to (\mathcal{V},\omega)$ *bilden eine Untergruppe der speziellen linearen Gruppe* $Sl(2L,\mathbb{C})$[1]. *Diese Gruppe bezeichnet man mit* $\mathfrak{S}\mathfrak{P}(2L,\mathbb{C})$.

Denn (vgl. [deG97, Kap. 1, S. 18/19]): Seien wiederum ν_1, ν_2 in $(\mathcal{V},\omega)$. Natürlich erhält die Identität die hermitesch symplektische Struktur (da $\omega\big(\mathbb{1}(\nu_1),\mathbb{1}(\nu_2)\big) = \omega(\nu_1,\nu_2)$). Auch die Komposition zweier hermitesch symplektischer Transformationen ist wieder hermitesch symplektisch, denn für zwei hermitesch symplektische Automorphismen s und t gilt: $\omega\big(s\circ t(\nu_1), s\circ t(\nu_2)\big) = s^*\omega\big(t(\nu_1),t(\nu_2)\big) = s^*\omega(\nu_1,\nu_2) = \omega(\nu_1,\nu_2)$. Mit s ist auch s^{-1} hermitesch symplektisch, denn: Falls $\omega\big(s(\nu_1),s(\nu_2)\big)$
$= \omega(\nu_1,\nu_2)$ ist, ersetze man ν_i durch $s^{-1}(\nu_i)$, $(i = 1,2)$. Damit erhält man: $\omega(\nu_1,\nu_2) = \omega\big(s^{-1}(\nu_1), s^{-1}(\nu_2)\big)$, womit auch die letzte Eigenschaft bewiesen ist.
□

Definition A.5 *Seien* $(\mathcal{V}_1,\omega_1)$, $(\mathcal{V}_2,\omega_2)$ *zwei hermitesch symplektische Räume, die beide die Dimension* $2L$ *haben, seien* ν_1^1, $\nu_2^1 \in (\mathcal{V}_1,\omega_1)$ *und* ν_1^2, $\nu_2^2 \in (\mathcal{V}_2,\omega_2)$. *Eine lineare Abbildung* $f: \mathcal{V}_1 \to \mathcal{V}_2$ *ist ein* hermitesch symplektischer Isomorphismus, *falls:*

$$\begin{aligned} & f^*\omega_1 = \omega_2 \\ \Leftrightarrow \quad & \omega_1(\nu_1^1,\nu_2^1) = \omega_2(f(\nu_1^2), f(\nu_2^2)). \end{aligned}$$

Satz A.2 *Man kann zeigen, dass zwischen zwei hermitesch symplektischen Räumen* $(\mathcal{V}_1,\omega_1)$ *und* $(\mathcal{V}_2,\omega_2)$ *der gleichen Dimension immer ein hermitesch symplektischer Isomorphismus* $f: \mathcal{V}_1 \to \mathcal{V}_2$ *existiert.*

Zum detaillierten **Beweis** siehe [deG97, Kap. 1,S. 16–19], hier wird nur kurz die Vorgehensweise skizziert: Der Isomorphismus f wird zunächst auf den beiden Basen der hermitesch symplektischen Vektorräume[2] erklärt. Ist $\{d_1^1,\ldots,d_L^1\} \cup \{e_1^1,\ldots,e_L^1\}$ eine Basis von (V_1,w_1) und $\{d_1^2,\ldots,d_L^2\} \cup \{e_1^2,\ldots,e_L^2\}$ eine Basis von

[1] Diese Gruppe umfasst die komplexen $2L\times 2L$ -Matrizen, deren Determinante gleich 1 ist.

[2] Die Basis eines hermitesch symplektischen Vektorraums $\mathcal{V}$ mit $\dim(\mathcal{V}) = 2L$ kann durch eine Menge von Vektoren $\{d_1,\ldots,d_L\} \cup \{e_1,\ldots,e_L\}$ erklärt werden, die folgende Voraussetzungen erfüllen: $\omega(d_i,d_j) = \omega(e_i,e_j) = 0$ für alle $i \neq j,\ 1 \le i,j \le L$ und $\omega(d_i,e_j) = \delta_{ij},\ 1 \le i,j \le L$. Dabei bezeichnet δ_{ij} das Kronecker-Symbol. Eine solche Basis existiert für jeden hermitesch symplektischen Raum (vgl. [deG97, Kap. 1, S. 16/17]).

(V_2, w_2), so wird der gesuchte Isomorphismus f durch $f(d_i^1) = d_i^2$ und $f(e_i^1) = e_i^2$ erklärt.

Dieses Resultat bedeutet insbesondere, dass alle hermitesch symplektischen Vektorräume (V_i, w_i) der Dimension $2L$ isomorph zu $\mathfrak{SP}(2L, \mathbb{C})$ sind. Es reicht also aus, Betrachtungen für die hermitesch symplektische Standardform – eine spezielle Darstellung von $\mathfrak{SP}(2L, \mathbb{C})$, die ich gleich definieren werde – anzustellen. Darüber hinaus besagt der *Satz von Darboux,* dass es sogar für jede endlichdimensionale Mannigfaltigkeit M mit einer hermitesch symplektischen Struktur ω eine Umgebung U und eine glatte Abbildung $f : U \to \mathbb{R}^{2L}$ gibt, für die gilt:

$$f^*\omega_0 = \omega.$$

Dabei ist ω_0 wiederum die hermitesch symplektische Standardform. Ein Beweis dazu findet sich etwa in [Be98, Kap. 2, S. 34–41].

Nun, endlich, stelle ich die schon mehrfach erwähnte Standardform vor:

Beispiel A.1 (Die hermitesch symplektische Standardform)
Das wichtigste Beispiel für einen endlichdimensionalen hermitesch symplektischen Raum ist der hermitesch symplektische Standardraum $(\mathbb{K}^{2L}, \omega_0)$, *wobei* ω_0 *für die* hermitesch symplektische Standardform *steht und* $\mathbb{K}$ *wie üblich einen Zahlenkörper (entweder* $\mathbb{R}$ *oder* $\mathbb{C}$*) symbolisiert.*
Die hermitesch symplektische Standardform definiert man wie folgt: Sei

$$\nu_1 = (\alpha_1,\ \ldots,\ \alpha_L,\ \beta_1,\ \ldots,\ \beta_L) \text{und} \nu_2 = (\widetilde{\alpha_1},\ \ldots,\ \widetilde{\alpha_L},\ \widetilde{\beta_1},\ \ldots,\ \widetilde{\beta_L}).$$

Dann ist:

$$\omega_0(\nu_1, \nu_2) = \sum_{j=1}^{L} \alpha_j \widetilde{\beta_j} - \widetilde{\alpha_j} \beta_j.$$

Insbesondere ergibt sich für $L = 1$*:* $\omega_0(\nu_1, \nu_2) = \det(\nu_1, \nu_2)$. *Im allgemeinen Fall beschreibt* $\omega_0(\nu_1, \nu_2)$ *die Summe der Flächen des Parallelogramms, das durch die Projektionen von* ν_1 *und* ν_2 *auf die Koordinatenebenen* α_j, β_j *aufgespannt wird. Daher bezeichnet man* ω_0 *auch oft als* Volumenform.
Die Gruppe aller Automorphismen s *von* $(\mathbb{C}^{2L}, \omega_0)$, *für die gilt:*

$$\omega_0\big(s(\nu_1), s(\nu_2)\big) = \omega_0(\nu_1, \nu_2) \qquad \text{für alle } \nu_1, \nu_2 \in \mathbb{C}^{2L},$$

heißt hermitesch symplektische Standardgruppe *(Notation:* $\mathfrak{SP}(2L, \mathbb{C})$*)*[3].

[3]Es ist auch möglich, die symplektische Standardform über $\mathbb{R}$ oder über die Quaternionen, $\mathbb{H}$, zu definieren. Um deutlich zu machen, über welchem Zahlenkörper (bzw. Schiefkörper) die symplektische Standardgruppe gebildet wird, schreibt man auch $\mathfrak{SP}(2L, \mathbb{K})$. Außerdem findet man in der Literatur oft auch die Bezeichnung $\mathfrak{SP}(L, \mathbb{K})$ (L statt $2L$).

Allgemeiner kann man auch hermitesch symplektische Mannigfaltigkeiten (M, ω) definieren (vgl. [Be98, Kap. 2, S. 33/34] sowie [deG97, Kap. 1, S. 28–34]), dabei ist M eine differenzierbare Mannigfaltigkeit und ω eine nichtdegenerierte, geschlossene 2-Form auf M. Eine differenzierbare 2-Form auf einem Vektorraum $\mathbb{C}^{2L}$ ist dadurch gegeben, dass man jedem ν in $\mathbb{C}^{2L}$ die Linearkombination

$$\zeta_\nu = \sum_{i<j\leq 2L} b_{ij}(x)\, dx_i \wedge dx_j$$

zuordnet. Dabei werden die b_{ij} gewöhnlich als C^∞-Funktionen gewählt und das Produkt $dx_i \wedge dx_j$ ist gegeben durch

$$dx_i \wedge dx_j = dx_i \otimes dx_j - dx_j \otimes dx_i,$$

die Abbildungs $dx_i : \mathbb{C}^{2L} \to \mathbb{C}$ ist die Projektion auf die i-te Koordinate. Insgesamt ergibt sich für $\nu_1,\ \nu_2$ in $\mathbb{C}^{2L}$:

$$d\alpha_j \wedge d\beta_j(\nu_1, \nu_2) = \alpha_j \widetilde{\beta}_j - \widetilde{\alpha}_j \beta_j.$$

Die hermitesch symplektische Standardform kann also mit $d\alpha \wedge d\beta = \sum_{j=1}^{L} d\alpha_j \wedge d\beta_j$ identifiziert werden. Aus Satz A.2 folgt damit sofort, dass die hermitesch symplektische Standardgruppe $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ isomorph zu der hermitesch symplektischen Gruppe $\mathfrak{S}\mathfrak{P}(\mathcal{V}, \omega)$ für jeden $2L$-dimensionalen hermitesch symplektischen Vektorraum $\mathcal{V}$ ist. Im folgenden werde ich daher alle meine Betrachtungen auf $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ beschränken.

Nun fragt man sich natürlich, wie sich die hermitesch symplektische Standardgruppe am sinnvollsten darstellen lässt. Besonders gut kann man $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ aufgefasst als Gruppe von Matrizen handhaben. Indem man eine geeignete Basis von $(\mathcal{V}, \omega)$ auswählt, identifiziert man $\mathfrak{M} \in \mathfrak{S}\mathfrak{P}$ mit seiner Matrix im hermitesch symplektischen Raum, d. h.

$$\mathfrak{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

A, B, C und D sind dabei in $\mathrm{Mat}(L \times L, \mathbb{C})$.

Die Anwendung dieser Konvention führt dazu, dass die hermitesch symplektische Standardform ω_0, zu der alle anderen hermitesch symplektischen Formen äquivalent sind, wie folgt geschrieben werden kann. Vorher sei noch einmal sei betont, dass dies nur dann möglich ist, wenn man die kanonische hermitesch symplektische Basis wählt, die auf die hermitesch symplektische Standardform führt. Das Aussehen dieser Standardform wurde schon im Beweis zu Lemma A.1 angedeutet:

Definition A.6 *Die hermitesch symplektische Standardform ergibt sich durch:*

$$\omega_0(\nu_1, \nu_2) = \nu_1^* \mathcal{J} \nu_2$$

mit Hilfe der Matrix

$$\mathcal{J} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} .$$

Eine Matrix $\mathfrak{M}$ *erhält die Form* (3.1) *also genau dann, wenn gilt:*

$$\mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J} .$$

Das heißt:

$$\mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) = \{\mathfrak{M} \in \; Mat(2L, \mathbb{C}) : \; \mathfrak{M}^* \mathcal{J} \mathfrak{M} = \mathcal{J}\} .$$

Anhang B

Verallgemeinerung der symplektischen Strukturen auf $\mathbb{R}$ und $\mathbb{H}$

Im dritten Kapitel wurde stets die hermitesch symplektische Gruppe $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ bzw. die hermitesch symplektische Lie-Algebra $\mathfrak{S}\mathfrak{P}(2L, \mathbb{C})$ betrachtet. Nun stellt man sich die Frage, wie die symplektischen Strukturen für Vektorräumen über $\mathbb{R}$ und $\mathbb{H}$ aussehen. Der zweite Teil des Anhangs ist genau dieser Verallgemeinerung gewidmet: Mit Hilfe von Theorem 1 lässt sich zeigen, dass die symplektische Struktur auf $\mathbb{R}$ und $\mathbb{H}$ übertragen werden kann, indem man zusätzliche Anforderungen an die Symmetrie der Matrizen stellt. Damit ist es möglich, auch die Lorentzgruppe sowie die Ergebnisse über die oberen Halbebenen und Einheitskreise verallgemeinern.

Die Frage, wie man die bisherigen Ergebnisse auf allgemeinere Vektorräume über $\mathbb{R}$ und $\mathbb{H}$ übertragen kann, wird in dieser Arbeit ähnlich wie in [SB06] erörtert. Besonders zu betonen ist, dass sich alle bisher definierten Strukturen, die auf der hermitesch symplektischen Gruppe beruhen, dadurch auf Vektorräume über $\mathbb{R}$ und $\mathbb{H}$ erweitern lassen, indem man schlicht zusätzliche Symmetrieeigenschaften fordert.

Vorweg möchte ich zunächst einige Bemerkungen zur Menge der Quaternionen schicken. Erinnert man sich daran, dass man $\mathbb{C}$ als reellen Vektorraum, der durch den Spann von i und 1 definiert wird, auffassen kann, ist folgende Definition der Quaternionen sehr leicht verständlich:

Definition B.1 *Eine Basis der Menge der Quaternionen*[1] $\mathbb{H}$ *(aufgefasst als Vektorraum über* $\mathbb{R}$*) ist gegeben durch* 1 *und die imaginären Einheiten* i, j *und* k, *die die Hamilton-Regeln*

$$i^2 = j^2 = k^2 = i\,j\,k = -1$$

befolgen (vgl. [Bo04, Kap. 2, S. 29]).

Im folgenden benötigt man neben $\mathcal{J}$ noch zwei weitere Matrizen, um die symplektischen Strukturen der Quaternionen zu charakterisieren. Man setze

$$I = \begin{pmatrix} 0 & -\mathbb{1}_{L/2} \\ \mathbb{1}_{L/2} & 0 \end{pmatrix}$$

und

$$\mathcal{I} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \mathbb{1}_{L/2} \\ 0 & 0 & -\mathbb{1}_{L/2} & 0 \\ 0 & -\mathbb{1}_{L/2} & 0 & 0 \\ \mathbb{1}_{L/2} & 0 & 0 & 0 \end{pmatrix}.$$

B.1 Lagrange-Ebenen und symplektische Gruppe über $\mathbb{R}$ und $\mathbb{H}$

Somit ergibt sich folgende Definition der reellen und quaternionischen Lagrange-Ebenen:

Definition B.2 (Lagrange-Ebenen)
Es sei

$$\begin{aligned} \widehat{\mathbb{L}}_L^{\mathbb{R}} = \big\{[\Phi]_\sim :\ \Phi \in \mathrm{Mat}\,(2L \times L,\, \mathbb{C}),& \\ \mathrm{Rang}\ (\Phi) = L,\ \Phi^* \mathcal{J} \Phi = 0,\ \Phi^T \mathcal{J} \Phi = 0\big\}& \end{aligned} \tag{B.1}$$

sowie

$$\begin{aligned} \widehat{\mathbb{L}}_L^{\mathbb{H}} = \big\{[\Phi]_\sim :\ \Phi \in \mathrm{Mat}\,(2L \times L,\, \mathbb{C}),& \\ \mathrm{Rang}\ (\Phi) = L,\ \Phi^* \mathcal{J} \Phi = 0,\ \Phi^T \mathcal{I} \Phi = 0\big\}.& \end{aligned} \tag{B.2}$$

[1]Da Sir William Rowan Hamilton (1805–1865) diese Definition der Quaternionen 1843 als erster benutzte, werden Quaternionen auch als *Hamilton-Zahlen* bezeichnet; damit ist die Symbolik, $\mathbb{H}$, erklärt. Quaternionen sind besonders gut dazu geeignet, dreidimensionale Räume und insbesondere Drehungen zu beschreiben. Sie bilden einen Schiefkörper.

Folgerung B.1 *Aus dieser Definition ergibt sich sofort:*

$$\widehat{\mathbb{L}}_L^{\mathbb{R}} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : \ a,\ b \ \in \ \mathrm{Mat}(L, \mathbb{C}), \right.$$
$$\left. \mathrm{Rang} \begin{pmatrix} a \\ b \end{pmatrix} = L,\ a^*a = b^*b,\ a^Tb = b^Ta \right\} \tag{B.3}$$

und

$$\widehat{\mathbb{L}}_L^{\mathbb{H}} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : \ a,\ b \ \in \ \mathrm{Mat}(L, \mathbb{C}), \right.$$
$$\left. \mathrm{Rang} \begin{pmatrix} a \\ b \end{pmatrix} = L,\ a^*a = b^*b,\ a^TIb = b^TIa \right\}. \tag{B.4}$$

Denn: Die Behauptung $a^*a = b^*b$ ergibt sich sofort aus (3.6). Den zweiten Teil erhält man, mit Hilfe der Bedingung $\Phi^T \mathcal{J} \Phi = 0$:

$$0 = \Phi^T \mathcal{J} \Phi = \begin{pmatrix} a^T & b^T \end{pmatrix} \begin{pmatrix} -b \\ a \end{pmatrix} = -a^Tb + b^Ta.$$

Auch für $\widehat{\mathbb{L}}_L^{\mathbb{H}}$ ergibt sich der erste Teil der Behauptung aus (3.6), den zweiten Teil rechnet man sofort nach:

$$0 = \Phi^T \mathcal{I} \Phi = \begin{pmatrix} a^T & b^T \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = -a^TIb + b^TIa. \qquad \square$$

Bemerkung B.1 *Beachtet man, dass der Repräsentant Φ einer Lagrange-Ebene $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{R}}$ stets so gewählt werden kann, dass $\Phi^*\Phi = \mathbb{1}$ ist, so ergibt sich:*

$$(\Phi,\ \mathcal{J}\Phi) \ \in \ \mathfrak{S}\mathfrak{P}(2L, \mathbb{C}) \cap \mathfrak{O}(2L) \cong \mathfrak{U}(L).$$

Die letzte Isomorphie wird in Satz 3.1 bewiesen. Verschiedene Orthonormalsysteme erhält man durch orthogonale Basiswechsel, die innerhalb der Lagrange-Ebene $[\Phi]_\sim$ bleiben. Also gilt:

$$\mathbb{L}_L^{\mathbb{R}} \cong \mathfrak{U}(L) \,/\, \mathfrak{O}(L).$$

Nun kann man auf eine recht einfache Art und Weise zeigen, dass es sich bei $\mathbb{L}_L^{\mathbb{R}}$ tatsächlich um eine kompakte Manngifaltigkeit handelt (für $\mathbb{L}_L^{\mathbb{C}}$ ist dies um einiges komplizierter und wurde im Laufe vierten Kapitel bewiesen).

Lemma B.1 *Mit Hilfe der Identifikation $\mathbb{L}_L^{\mathbb{R}} \cong \mathfrak{U}(L) \,/\, \mathfrak{O}(L)$ sieht man, dass $\mathbb{L}_L^{\mathbb{R}}$ die Struktur einer kompakten analytischen Mannigfaltigkeit hat.*

Beweis: Ähnliche Beweise finden sich auch häufig in der Literatur (vgl. beispielsweise [Ar67, CGIP03]).

Sei $[U]$ eine der obigen Rechtsnebenklassen, d. h. $[U] = U\,O$, wobei $U \in \mathfrak{U}(L)$ und $O \in \mathfrak{O}(L)$ ist. Dann gilt:

$$UO\,(UO)^T = U\,\underbrace{OO^T}_{=\mathbb{1}_L}\,U^T = UU^T =: V.$$

Dann ist $V \in \mathfrak{U}(L) \cap \mathrm{Sym}(L, \mathbb{C})$, da mit U auch U^T unitär ist und da $(UU^T)^T = UU^T$ ist.

Andererseits ist jedes $V \in \mathfrak{U}(L) \cap \mathrm{Sym}(L, \mathbb{C})$ orthogonal diagonalisierbar. Somit hat V eine Darstellung als $V = O\,D\,O^T$. Dabei ist $O \in \mathfrak{O}(L)$; $D \in \mathrm{Diag}(L, \mathbb{R})$ ist eine Diagonalmatrix. Daher existiert die Wurzel von D, bezeichnet mit $D^{\frac{1}{2}}$.[2] Daraus folgt:

$$V = O\,D\,O^T = OD^{\frac{1}{2}}\,D^{\frac{1}{2}}O^T = \left(OD^{\frac{1}{2}}\right)\,\left(OD^{\frac{1}{2}}\right)^T.$$

Nun setzt man $U = OD^{\frac{1}{2}}$. Damit folgt sofort die Behauptung. □

Bemerkung B.2 *Randbedingungen Im komplexen Fall wurden stets Randbedingungen der Form $\begin{pmatrix} y_0^* & (y')_0^* \end{pmatrix}^* \Phi_\xi = \begin{pmatrix} \xi^* & \mathbb{1}^* \end{pmatrix}^*$ betrachtet (analog für $x = 1$). Dabei wurde gefordert, dass $\xi \in \mathrm{Her}(L, \mathbb{C})$ sei, da Φ_ξ dann eine Lagrange-Ebene ist. Möchte man diese Vorgehensweise auf $\mathbb{R}$ und $\mathbb{R}$ übertragen, so fordert man im Falle $\mathbb{K} = \mathbb{R}$:*

$$\xi \in \mathrm{Sym}(L, \mathbb{R})$$

und im Falle $\mathbb{K} = \mathbb{H}$:

$$\xi \in \mathrm{Sym}(L, \mathbb{R}) \cap \mathrm{Self}(L, \mathbb{C}).$$

[2] Diese Wurzel bildet man einfach, indem man die Wurzeln der Einträge auf der Hauptdiagonalen bildet. Offensichtlich gilt dann: $D^{\frac{1}{2}}D^{\frac{1}{2}} = D$. Zudem ist, wie für alle Diagonalmatrizen, $(D^{\frac{1}{2}})^T = D^{\frac{1}{2}}$.

Mit Definition B.2 erhält man die folgende Erweiterung von Theorem 1:

Satz B.1

1. *Die Abbildung* $\Pi : \widehat{\mathbb{L}}_L^{\mathbb{R}} \to U(L) \cap \mathrm{Sym}\,(L, \mathbb{C})$ *ein reeller analytischer Diffeomorphismus.*

2. *Ist* L *gerade, so ist die Abbildung* $\Pi : \widehat{\mathbb{L}}_L^{\mathbb{H}} \to U(L) \cap \mathrm{Self}\,(L, \mathbb{C})$ *ist ein reeller analytischer Diffeomorphismus.*

Beweis (Beweisidee nach [SB06]): In beiden Fällen sei stets $\Phi = \begin{pmatrix} a \\ b \end{pmatrix}$.

1. Wie im Beweis von Theorem 1 muss man im ersten Beweisschritt zeigen, dass die Cayley-Transformation von $\widehat{\mathbb{L}}_L^{\mathbb{R}}$ die folgende Gestalt hat:

$$\mathcal{C}\,\widehat{\mathbb{L}}_L^{\mathbb{R}} = \left\{ [\Phi]_\sim : \ a,\ b \in Gl(L, \mathbb{C}) : a^* a = b^* b \text{ und } (ab^{-1})^T = ab^{-1} \right\}.$$

Dies gilt, denn: $\mathcal{C}\,\Phi = \begin{pmatrix} a - ib \\ a + ib \end{pmatrix} =: \begin{pmatrix} \widehat{a} \\ \widehat{b} \end{pmatrix}$. Daraus folgt nach dem Beweis von Theorem 1, dass $\widehat{a}^*\,\widehat{a} = \widehat{b}^*\,\widehat{b}$ ist. Auch die Invertierbarkeit von $\widehat{a}$ und $\widehat{b}$ folgt sofort aus dem Beweis von Theorem 1. Beachtet man nun, dass nach (B.3) $(b^T)^{-1} a^T = b^{-1} a$ und $(a^T)^{-1} b^T = b a^{-1}$ ist, so folgt:

$$\begin{aligned}
(\widehat{a}\,\widehat{b}^{-1})^T &= \big((a - ib)(a + ib)^{-1}\big)^T \\
&= \big((a - ib)(a^{-1} + \frac{1}{i} b^{-1}\big)^T \\
&= \big(aa^{-1} + \frac{1}{i} ab^{-1} - iba^{-1} - bb^{-1}\big)^T \\
&= \mathbb{1} - i(b^{-1})^T a^T - i(a^{-1})^T b^T - \mathbb{1} \\
&= -ib^{-1} a - iba^{-1} \\
&= (a - ib)(a + ib)^{-1} = \widehat{a}\,\widehat{b}^{-1}.
\end{aligned}$$

Das impliziert, dass $\Pi\,([\Phi]_\sim) = \widehat{a}\,\widehat{b}^{-1}$ für $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{R}}$ symmetrisch ist. Die Inverse von Π ist nach (4.3): $\Pi^{-1}(U) = \left[\begin{pmatrix} \frac{1}{2}(U + \mathbb{1}) \\ \frac{i}{2}(U - \mathbb{1}) \end{pmatrix}\right]_\sim$. Falls U dabei symmetrisch ist, gilt: $\big(\Pi^{-1}(U)\big)^T \mathcal{J}\,\Pi^{-1}(U) = 0$, denn:

$$\begin{aligned}
\big(\Pi^{-1}(U)\big)^T \mathcal{J}\,\Pi^{-1}(U) &= \begin{pmatrix} \frac{1}{2}(U+\mathbb{1}) \\ \frac{i}{2}(U-\mathbb{1}) \end{pmatrix}^T \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(U+\mathbb{1}) \\ \frac{i}{2}(U-\mathbb{1}) \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}(U^T+\mathbb{1}) & \frac{i}{2}(U^T-\mathbb{1}) \end{pmatrix} \begin{pmatrix} -\frac{i}{2}(U-\mathbb{1}) \\ \frac{1}{2}(U+\mathbb{1}) \end{pmatrix} \\
&= -\frac{i}{4}(\underbrace{U^T}_{=U}+\mathbb{1})(U-\mathbb{1}) + \frac{i}{4}(\underbrace{U^T}_{=U}-\mathbb{1})(U+\mathbb{1}) \\
&= 0.
\end{aligned}$$

Im nächsten Schritt hat man noch Stetigkeit und reelle Analytizität zu zeigen, diese folgen aber wie im Beweis zu Theorem 1, da auch in diesem Falle Π stetig und Π^{-1} reell analytisch ist.

2. Auch für die Quaternionen bildet man zunächst die Cayley-Transformierte der Lagrange-Ebenen:

$$\mathcal{C}\,\mathbb{L}_L^{\mathbb{H}} = \left\{ \left[\begin{pmatrix} a \\ b \end{pmatrix}\right]_\sim \; : \; a,\ b \,\in\, Gl(L,\mathbb{C}),\ a^*a = b^*b,\ I^*(ab^{-1})^T I = ab^{-1} \right\}.$$

Setzt man wie schon im ersten Teil $\mathcal{C}\,\Phi = \begin{pmatrix} \widehat{a} \\ \widehat{b} \end{pmatrix}$ so folgt wiederum $\widehat{a}^*\,\widehat{a} = \widehat{b}^*\,\widehat{b}$ und die Invertierbarkeit von $\widehat{a}$ und $\widehat{b}$ genauso wie im Beweis von Theorem 1. Um die zusätzlich geforderte Symmetrieeigenschaft nachzuweisen, benutzt man (B.4). Danach ist $a^T I b = b^T I a$. Sind a und b invertierbar, ist das äquivalent zu $I\,b\,a^{-1} = (a^T)^{-1} b^T I$ und $(b^T)^{-1} a^T I = I\,a\,b^{-1}$. Beachtet man, dass $I^* = -I = I^{-1}$ ist, folgt $I^*(a^T)^{-1} b^T I = b\,a^{-1}$ und $I^*(b^T)^{-1} a^T I = a\,b^{-1}$. Damit rechnet man $I^*(\widehat{a}\,\widehat{b}^{-1})^T I$ aus:

$$\begin{aligned}
I^*(\widehat{a}\,\widehat{b}^{-1})^T I &= I^*\left((a-ib)(a+ib)^{-1}\right)^T I \\
&= I^*\left((a-ib)(a^{-1}-ib^{-1})\right)^T I \\
&= I^*\big(-i(b^{-1})^T a^T - i(a^{-1})^T b^T\big)\, I \\
&= -i\, I^*(b^{-1})^T a^T\, I - i\, I^*(a^{-1})^T b^T\, I \\
&= -i\,a\,b^{-1} - i\,b\,a^{-1} \\
&= (a - i\,b)(a^{-1} - i\,b^{-1}) = \widehat{a}\,\widehat{b}^{-1}.
\end{aligned}$$

Das bedeutet, $I\,\Pi\left([\Phi]_\sim\right)$ ist für $[\Phi]_\sim \in \mathbb{L}_L^{\mathbb{H}}$ schiefsymmetrisch, denn

$$\begin{aligned} I\,\Pi\left([\Phi]_\sim\right) &= I\,\widehat{a}\,\widehat{b}^{-1} \\ &= I\,I^* \left(\widehat{a}\,\widehat{b}^{-1}\right)^T I \\ &= \mathbb{1}_L \left(\widehat{a}\,\widehat{b}^{-1}\right)^T I \\ &= -\left(\widehat{a}\,\widehat{b}^{-1}\right)^T I^T = -\big(I\,\Pi\left([\Phi]_\sim\right)\big)^T. \end{aligned}$$

Falls $I\,U$ ebenfalls schiefsymmetrisch ist, folgt:

$$I\,U = -(I\,U)^T = -U^T I^T = U^T I.$$

Somit gilt nach (4.3):

$$\begin{aligned} \left(\Pi^{-1}(U)\right)^T \mathcal{I}\,\Pi^{-1}(U) &= \begin{pmatrix} \frac{1}{2}(U+\mathbb{1}) \\ \frac{i}{2}(U-\mathbb{1}) \end{pmatrix}^T \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(U+\mathbb{1}) \\ \frac{i}{2}(U-\mathbb{1}) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2}(U^T+\mathbb{1}) & \frac{i}{2}(U^T-\mathbb{1}) \end{pmatrix} \begin{pmatrix} -\frac{i}{2}I(U-\mathbb{1}) \\ \frac{1}{2}I(U+\mathbb{1}) \end{pmatrix} \\ &= -\frac{i}{4}\Big(U^T\,I\,U \underbrace{-U^T\,I + I\,U}_{=0} - I\Big) \\ &\quad +\frac{i}{4}\Big(U^T\,I\,U + \underbrace{U^T\,I - I\,U}_{=0} - I\Big) \\ &= 0. \end{aligned}$$

Für schiefsymmetrische $I\,U$ ist $\Pi^{-1}(U)$ also in $\widehat{\mathbb{L}}_L^{\mathbb{H}}$. Berücksichtigt man noch die Stetigkeit von Π und die Analytizität von Π^{-1}, ist der Beweis vollständig. □

Mit Hilfe dieser Ergebnisse lassen sich nun die symplektischen Gruppen über $\mathbb{R}$ und $\mathbb{H}$ definieren. Doch zunächst eine weitere Definition:

Definition B.3 *Sei $\mathcal{A}$ eine Matrix mit reellen Einträgen, dann definiert man:*

$$\mathcal{A}^{*\mathbb{R}} = \mathcal{T}^T.$$

Wenn $\mathcal{B}$ dagegen eine Matrix mit Einträgen aus dem Schiefkörper der Quaternionen ist, so sei:

$$\mathcal{B}^{*\mathbb{H}} = \overline{\mathcal{B}^{\,T}}^{\,\mathbb{H}},$$

dabei bezeichne $\overline{\mathcal{B}}^{\,\mathbb{H}}$ die quaternionisch konjugierte Matrix zu $\mathcal{B}^{\mathbb{H}}$. Analog zur komplexen Konjugation werden dabei die Vorzeichen der drei imaginären Einheiten i, j und k umgekehrt.

Schließlich definiert man auf Basis obiger Definition die reelle und die quaternionische symplektische Gruppe:

Definition B.4 *Die reelle symplektische Gruppe ist gegeben durch*

$$\mathfrak{SP}(2L, \mathbb{R}) = \left\{ \mathcal{T} \in \text{Mat}\,(2L, \mathbb{R}) : \; \mathcal{T}^{*\mathbb{R}} \mathcal{J} \mathcal{T} = \mathcal{J} \right\}.$$

Ist darüber hinaus L gerade, so ist die quaternionische symplektische Gruppe definiert als

$$\mathfrak{SP}(2L, \mathbb{H}) \cong \left\{ \mathcal{T} \in \text{Mat}\,(L, \mathbb{H}) : \; \mathcal{T}^{*\mathbb{H}} I \mathcal{T} = I \right\}.$$

Diese Definition ist sinnvoll, da man zeigen kann, dass die Lagrange-Ebenen $\mathbb{L}_L^{\mathbb{R}}$ bzw. $\mathbb{L}_L^{\mathbb{H}}$ unter Transformation mit symplektischen Matrizen aus $\mathfrak{SP}(2L, \mathbb{R})$ bzw. $\mathfrak{SP}(2L, \mathbb{H})$ invariant sind:

Satz B.2 *Im Sinne von diffeomorphen reell analytischen Mannigfaltigkeiten gilt:*

$$\begin{aligned} \mathbb{L}_L^{\mathbb{R}} &= \left\{ [\Phi]_\sim : \; \Phi \in \text{Mat}\,(2L \times L, \; \mathbb{R}), \; \text{Rang}\; (\Phi) = L, \; \Phi^{*\mathbb{R}} \mathcal{J} \Phi = 0 \right\} \\ &= \widehat{\mathbb{L}}_L^{\mathbb{R}}. \end{aligned}$$

Außerdem folgt:

$$\begin{aligned} \mathbb{L}_L^{\mathbb{H}} &= \left\{ [\Phi]_\sim : \; \Phi \in \text{Mat}\,(L \times L, \; \mathbb{H}), \; \text{Rang}\; (\Phi) = L, \; \Phi^{*\mathbb{H}} I \Phi = 0 \right\} \\ &= \widehat{\mathbb{L}}_L^{\mathbb{H}}. \end{aligned}$$

Beweis (Idee nach [SB06]): Es genügt zu zeigen, dass die beiden jeweils angegebenen Definitionen $\mathbb{L}_L^{\mathbb{R}}$ und $\widehat{\mathbb{L}}_L^{\mathbb{R}}$ sowie $\mathbb{L}_L^{\mathbb{H}}$ und $\widehat{\mathbb{L}}_L^{\mathbb{H}}$ übereinstimmen.

Zunächst zu $\mathbb{L}_L^{\mathbb{R}}$: Die Inklusion $\mathbb{L}_L^{\mathbb{R}} \subset \widehat{\mathbb{L}}_L^{\mathbb{R}}$ folgt daraus, dass die beiden Bedingungen $\Phi^* \mathcal{J} \Phi = 0$ und $\Phi^T \mathcal{J} \Phi = 0$ für reelle Φ offensichtlich äquivalent sind. Darüber hinaus ist die Inklusion stetig.

Für die umgekehrte Inklusion genügt es nach dem ersten Teil von Satz B.1 zu zeigen, dass gilt: $\Pi^{-1} : \mathfrak{U}(L, \mathbb{C}) \cap \text{Sym}\,(L, \mathbb{C}) \to \mathbb{L}_L^{\mathbb{R}}$. Aufgrund des Definitionsbereichs von Π^{-1} kann man für U aus (4.3) einen reellen Repräsentanten finden. Um dies zu beweisen, nutzt man aus, dass jede symmetrische unitäre Matrix U orthogonal diagonalisierbar ist. Das bedeutet, es existiert ein $M \in \mathfrak{O}(L)$, das $U = M^T\, D\, M$ erfüllt. Dabei ist $D = \text{diag}\big(e^{i\theta_1}, \; \ldots, \; e^{i\theta_L}\big)$ mit $\theta_l \in [0, \, 2\pi)$ für

$l = 1, \ldots, L$. Nun kann man die beiden Wurzeln bestimmen. Hier wählt man den positiven Zweig und setzt $D^{1/2} := \operatorname{diag}\left(e^{i\theta_1/2}, \ldots, e^{i\theta_L/2}\right)$. Damit ergibt sich:

$$V\,V^T = M^T\,D^{1/2}\,M\,\left(M^T\,D^{1/2}\,M\right)^T = M^T\,D^{1/2}\underbrace{M\,M^T}_{=\mathbb{1}}D^{1/2}M = M^T\,D\,M = U.$$

Nun setzt man $a = \Re(V)$ und $b = -\Im(V)$, dann gilt:

$$\begin{aligned}
\Pi^{-1}(U) &= \left[\begin{pmatrix} \frac{1}{2}(VV^T + \mathbb{1}) \\ \frac{i}{2}(VV^T - \mathbb{1}) \end{pmatrix}\right]_\sim = \left[\begin{pmatrix} \frac{1}{2}(VV^T + VV^*) \\ \frac{i}{2}(VV^T - VV^*) \end{pmatrix}\right]_\sim \\
&= \left[\begin{pmatrix} \frac{1}{2}\big((a-ib)(a-ib)^T + (a-ib)(a+ib)^T\big) \\ \frac{i}{2}\big((a-ib)(a-ib)^T - (a-ib)(a+ib)^T\big) \end{pmatrix}\right]_\sim \\
&= \left[\begin{pmatrix} \frac{1}{2}\big(aa^T - iab^T - iba^T - bb^T + aa^T - iba^T + iab^T + bb^T\big) \\ \frac{i}{2}\big(aa^T - iab^T - iba^T - bb^T - (aa^T - iba^T + iab^T + bb^T)\big) \end{pmatrix}\right]_\sim \\
&= \left[\begin{pmatrix} \frac{1}{2}\big(2aa^T - 2iba^T\big) \\ \frac{i}{2}\big(-2iab^T - 2bb^T\big) \end{pmatrix}\right]_\sim \\
&= \left[\begin{pmatrix} Va^T \\ Vb^T \end{pmatrix}\right]_\sim = \left[V\,\begin{pmatrix} a^T \\ b^T \end{pmatrix}\right]_\sim .
\end{aligned}$$

Das bedeutet, $\Phi = \begin{pmatrix} a^T \\ b^T \end{pmatrix}$. Außerdem ist Π^{-1} tatsächlich die Inverse von Π, denn es gilt:

$$\Pi\left([\Phi]_\sim\right) = \pi\begin{pmatrix} a^T - ib^T \\ a^T + ib^T \end{pmatrix} = \pi\begin{pmatrix} V^T \\ \overline{V}^T \end{pmatrix} = V^T(\overline{V}^T)^{-1} = V^T V = U.$$

Damit ist der erste Teil der Behauptung bewiesen.

$\mathbb{L}_L^{\mathbb{H}} = \widehat{\mathbb{L}}_L^{\mathbb{H}}$: Um dies zu zeigen, betrachtet man die Symmetrien der symplektischen Strukturen etwas näher. Eine Basis von $\mathbb{H}$ ist gegeben durch 1 sowie den imaginären Einheiten i, j und k. Nun definiert man die Matrix

$$\Upsilon = \begin{pmatrix} \mathbb{1}_{L/2} & j\mathbb{1}_{L/2} & 0 & 0 \\ 0 & 0 & \mathbb{1}_{L/2} & j\mathbb{1}_{L/2} \end{pmatrix} \in \operatorname{Mat}\left(L \times 2L,\, \mathbb{H}\right),$$

wobei alle Blocks von der Größe $2L \times 2L$ sind. Damit gilt:

$$\begin{aligned}
\Upsilon^{*\mathbb{H}}\, I\, \Upsilon &= \begin{pmatrix} \mathbb{1} & j\mathbb{1} & 0 & 0 \\ 0 & 0 & \mathbb{1} & j\mathbb{1} \end{pmatrix}^{*\mathbb{H}} \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \mathbb{1} & j\mathbb{1} & 0 & 0 \\ 0 & 0 & \mathbb{1} & j\mathbb{1} \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{1} & 0 \\ -j\mathbb{1} & 0 \\ 0 & \mathbb{1} \\ 0 & -j\mathbb{1} \end{pmatrix} \begin{pmatrix} 0 & 0 & -\mathbb{1} & -j\mathbb{1} \\ \mathbb{1} & j\mathbb{1} & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -\mathbb{1} & -j\mathbb{1} \\ 0 & 0 & j\mathbb{1} & -\mathbb{1} \\ \mathbb{1} & j\mathbb{1} & 0 & 0 \\ -j\mathbb{1} & \mathbb{1} & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -\mathbb{1} & 0 \\ 0 & 0 & 0 & -\mathbb{1} \\ \mathbb{1} & 0 & 0 & 0 \\ 0 & \mathbb{1} & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & j\mathbb{1} \\ 0 & 0 & -j\mathbb{1} & 0 \\ 0 & -j\mathbb{1} & 0 & 0 \\ j\mathbb{1} & 0 & 0 & 0 \end{pmatrix} \\
&= \mathcal{J} - j\,\mathcal{I}.
\end{aligned}$$

Das bedeutet, dass die Abbildung

$$\begin{aligned} \Upsilon: \ \mathrm{Mat}\,(2L \times L, \mathbb{C}) &\to \mathrm{Mat}\,(L \times L, \mathbb{H}) \\ A &\mapsto \Upsilon\, A \end{aligned}$$

eine Abbildung auf die (komplexe) Grassmann-Mannigfaltigkeit der Rechtsäquivalenzklassen ist. Durch Υ ist daher eine offenbar analytische Abbildung $\Upsilon : \mathbb{L}_L^{\mathbb{H}} \to \widehat{\mathbb{L}}_L^{\mathbb{H}}$ gegeben, und damit ist der Satz vollständig bewiesen. □

Bemerkung B.3

1. *Es sei festgehalten, dass offensichtlich für* $\mathcal{A} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{R})$ *und* $\mathcal{B} \in \mathfrak{S}\mathfrak{P}(2L, \mathbb{H})$ *sowie für zwei Lagrange-Ebenen* $\Phi_{\mathbb{R}} \in \widehat{\mathbb{L}}_L^{\mathbb{R}}$ *und* $\Phi_{\mathbb{H}} \in \widehat{\mathbb{L}}_L^{\mathbb{H}}$ *gilt:*

   $$\mathcal{A}\,\Phi_{\mathbb{R}} \in \widehat{\mathbb{L}}_L^{\mathbb{R}} \qquad \text{und} \qquad \mathcal{B}\,\Phi_{\mathbb{H}} \in \widehat{\mathbb{L}}_L^{\mathbb{H}},$$

   denn es gilt:

   $$(\mathcal{A}\,\Phi_{\mathbb{R}})^T\, \mathcal{J}\, \mathcal{A}\, \Phi_{\mathbb{R}} = \Phi_{\mathbb{R}}^T\, \underbrace{\mathcal{A}^T\, \mathcal{J}\, \mathcal{A}}_{=\mathcal{J}}\, \Phi_{\mathbb{R}} = \Phi_{\mathbb{R}}^T\, \mathcal{J}\, \Phi_{\mathbb{R}} = 0$$

sowie (vollkommen analog)

$$(\mathcal{B}\,\Phi_{\mathbb{H}})^T\ \mathcal{I}\,\mathcal{B}\,\Phi_{\mathbb{H}} = 0.$$

Dass die Bedingung $(\mathcal{A}\,\Phi_{\mathbb{R}})^*\ \mathcal{J}\,\mathcal{A}\,\Phi_{\mathbb{R}} = 0$ bzw. $(\mathcal{B}\,\Phi_{\mathbb{H}})^*\ \mathcal{J}\,\mathcal{B}\,\Phi_{\mathbb{H}} = 0$ erfüllt ist, ergibt sich genauso wie im Fall der hermitesch symplektischen Gruppe, den Beweis entnehme man Lemma 3.4.

2. *Möchte man die symplektischen Gruppen $\mathfrak{SP}(2L,\mathbb{R})$ und $\mathfrak{SP}(2L,\mathbb{H})$ in Blockmatrixform darstellen, ergibt sich für beide Gruppen wie für die hermitesch symplektische Gruppe mit $\mathbb{K} = \mathbb{C},\ \mathbb{R}$ oder $\mathbb{H}$:*

$$\mathfrak{SP}(2L,\mathbb{K}) = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \text{Mat}\ (2L,\mathbb{C}) : \right.$$
$$\left. A^*C = C^*A,\ B^*D = D^*B,\ A^*D - C^*B = \mathbb{1} \right\}.$$

Der Beweis dazu findet sich in Lemma 3.1. Darüber hinaus sind $\mathfrak{SP}(2L,\mathbb{R})$ und $\mathfrak{SP}(2L,\mathbb{H})$ durch zusätzliche Symmetrien charakterisiert. Ist $\mathcal{T} \in \mathfrak{SP}(2L,\mathbb{R})$, so sind die Blockmatrizen A, B, C und D reell. Wenn $\mathcal{T}$ dagegen in $\mathfrak{SP}(2L,\mathbb{H})$ ist, erhält man selbstkonjugierte[3] Blockmatrizen A, B, C und D.

B.2 Verallgemeinerung der Lorentzgruppe für $\mathbb{R}$ und $\mathbb{H}$

Jetzt kann man diese Ergebnisse sehr einfach auf die Lorentz-Gruppe, die oberen Halbebenen und die Einheitskreise übertragen, indem man für die reelle bzw. quaternionische Version zusätzliche Symmetrieeigenschaften fordert.

Definition B.5 *Die reelle Lorentzgruppe ist definiert durch:*

$$\mathcal{U}(L,L,\mathbb{R}) = \left\{\mathcal{T}\ \in\ \text{Mat}\ (2L \times 2L,\mathbb{C}) :\ \mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G},\ \mathcal{T}^T\mathcal{J}\mathcal{T} = \mathcal{J}\right\}.$$

Dagegen ist die Lorentzgruppe über die Quaternionen gegeben durch:

$$\mathcal{U}(L,L,\mathbb{H}) = \left\{\mathcal{T}\ \in\ \text{Mat}\ (2L \times 2L,\mathbb{C}) :\ \mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G},\ \mathcal{T}^T\mathcal{I}\mathcal{T} = \mathcal{I}\right\}.$$

[3] Das bedeutet, $\overline{A}^{\,\mathbb{H}} = A$.

Mit (4.4) ergibt sich, dass die verallgemeinerte Lorentzgruppe gegeben ist durch:

$$\mathcal{U}(L,L,\mathbb{K}) = \mathcal{C}\ \mathfrak{S}\mathfrak{P}(2L,\mathbb{K})\ \mathcal{C}^*.$$

Lemma B.2 (Blockmatrixdarstellung)

$$\begin{aligned}\mathcal{U}(L,L,\mathbb{K}) &= \Bigg\{\begin{pmatrix} A & B \\ C & D\end{pmatrix} \in \mathrm{Mat}(2L,\mathbb{C}) : \\ &\qquad A^*A - C^*C = \mathbb{1},\ D^*D - B^*B = \mathbb{1},\ A^*B = C^*D\Bigg\} \\ &= \Bigg\{\begin{pmatrix} A & B \\ C & D\end{pmatrix} \in \mathrm{Mat}(2L,\mathbb{C}) : \\ &\qquad AA^* - BB^* = \mathbb{1},\ DD^* - CC^* = \mathbb{1},\ AC^* = BD^*\Bigg\},\end{aligned}$$

wobei für $\mathbb{K} = \mathbb{R}$ zusätzlich gilt $C = \overline{B}$ und $A = \overline{D}$, wohingegen man für $\mathbb{K} = \mathbb{H}$ fordert, dass $C = I^*\ \overline{B}\ I$ und $A = I^*\ \overline{D}\ I$ ist.

Beweis: Die Behauptung für $\mathbb{K} = \mathbb{C}$ entspricht Lemma 4.3. Da die Lorentz-Gruppen über $\mathbb{R}$ und $\mathbb{H}$ nur durch *zusätzliche* Symmetrien gekennzeichnet sind, muss man nur noch die jeweiligen Behauptungen über die zusätzlichen Strukturen für $\mathbb{R}$ und $\mathbb{H}$ überprüfen:

Sei stets $\mathcal{T} = \begin{pmatrix} A & B \\ C & D\end{pmatrix}$. Dann gilt für $\mathcal{T} \in \mathcal{U}(L,L,\mathbb{R})$ (neben der bereits berücksichtigten Forderung $\mathcal{T}^*\mathcal{G}\mathcal{T} = \mathcal{G}$):

$$\mathcal{T}^T\mathcal{J}\mathcal{T} = \mathcal{J}.$$

Daraus folgt:

$$\begin{aligned}\mathcal{T}^T\mathcal{J}\mathcal{T} &= \begin{pmatrix} A^T & C^T \\ B^T & D^T\end{pmatrix}\begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0\end{pmatrix}\begin{pmatrix} A & B \\ C & D\end{pmatrix} \\ &= \begin{pmatrix} C^TA - A^TC & C^TB - A^TD \\ D^TA - B^TC & D^TB - B^TD\end{pmatrix} = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0\end{pmatrix} = \mathcal{J}.\end{aligned}$$

Nach (4.7) gilt aber, dass $D^*D - B^*B = \mathbb{1}$ ist, oder äquivalent dazu: $B^*B - D^*D = -\mathbb{1}$. Setzt man $C = \overline{B}$ und $A = \overline{D}$ so folgt:

$$C^TB - A^TD = \overline{B}^TB - \overline{D}^TD = B^*B - D^*D = -\mathbb{1}. \tag{B.5}$$

Darüber hinaus ergibt sich:

$$C^T A - A^T C = B^* A - D^* C \overset{(4.7)}{=} 0.$$

Weiterhin gilt:

$$(D^T A - B^T C)^T = A^T D - C^T B \overset{(B.5)}{=} \mathbb{1},$$

woraus man sofort schließt, dass auch $D^T A - B^T C = \mathbb{1}$ ist. Nun verifiziert man noch:

$$D^T B - B^T D = B^T D - D^T B = A^* B - C^* D \overset{(4.7)}{=} 0.$$

Damit sind alle Behauptungen für $\mathcal{U}(L, L, \mathbb{R})$ bewiesen.

Jetzt widmet man sich $\mathcal{U}(L, L, \mathbb{H})$: Die zusätzliche Symmetrie ist gegeben durch $\mathcal{T}^T \mathcal{I} \mathcal{T} = \mathcal{I}$. Mit der Formel für $\mathcal{T}$ erhält man:

$$\begin{aligned} \mathcal{T}^T \mathcal{I} \mathcal{T} = &= \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \\ &= \begin{pmatrix} C^T\, I\, A - A^T\, I\, C & C^T\, I\, B - A^T\, I\, D \\ D^T\, I\, A - B^T\, I\, C & D^T\, I\, B - B^T\, I\, D \end{pmatrix} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}. \end{aligned}$$

Setzt man nun wie in der Behauptung $C = I\, \overline{B}\, I^*$ sowie $A = I\, \overline{D}\, I^*$, und beachtet man darüber hinaus, dass $I^* = I^T = I^{-1}$ ist, so ergibt sich:

$$C^T IB - A^T ID = IB^* I^* IB - ID^* I^* ID = IB^* B - ID^* D = I(B^* B - D^* D) \overset{(4.7)}{=} -I.$$

Beachtet man, dass $I^T = -I$ ist, so ergibt sich:

$$\left(D^T IA - B^T IC\right)^T = -\left(C^T I^T B - A^T I^T D\right) = C^T IB - A^T ID = -I.$$

Daraus folgt:

$$D^T\, I\, A - B^T\, I\, C = -I^T = I.$$

Weiterhin ist

$$C^T\, I\, A - A^T\, I\, C = IB^* I^* IA - ID^* I^* IC = I(B^* A - D^* C) \overset{(4.7)}{=} 0.$$

Ebenso gilt, wenn man berücksichtigt, dass $II = \mathbb{1}$ ist:

$$D^T\, I\, B - B^T\, I\, D = I^* A^* IIB - I^* C^* IID = I^*(A^* B - C^* D) \overset{(4.7)}{=} 0.$$

Somit ist der Beweis komplett. □

B.3 Obere Halbebenene und Einheitskreis über $\mathbb{R}$ und $\mathbb{H}$

Auch die oberen Halbebenen und die Einheitskreise lassen sich nicht nur über $\mathbb{C}$, sondern – unter Zuhilfenahme weiterer Symmetrieforderungen – auch über $\mathbb{R}$ und $\mathbb{H}$ definieren:

Definition B.6 *Es seien $\mathbb{U}_L^{\mathbb{R}}$ und $\mathbb{U}_L^{\mathbb{H}}$ bzw. $\mathbb{D}_L^{\mathbb{R}}$ und $\mathbb{D}_L^{\mathbb{H}}$ die folgenden Untergruppen von $\mathbb{U}_L^{\mathbb{C}}$ bzw. $\mathbb{D}_L^{\mathbb{C}}$:*

$$\mathbb{U}_L^{\mathbb{R}} = \mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C}), \qquad \mathbb{D}_L^{\mathbb{R}} = \mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C})$$
$$\mathbb{U}_L^{\mathbb{H}} = \mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C}), \qquad \mathbb{D}_L^{\mathbb{H}} = \mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C}).$$

Weiterhin ist zu beachten, dass aus obiger Definition folgt:

$$I\mathbb{D}_L^{\mathbb{H}} = I\big(\mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C})\big) = \mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Skew}\,(L,\mathbb{C}).$$

Wie für $\mathbb{K} = \mathbb{C}$ gilt auch in diesem Falle:

Satz B.3 *Die Formeln*

$$U = (Z - i\mathbb{1})(Z + i\mathbb{1})^{-1} \qquad \text{und} \qquad Z = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1}$$

beschreiben einen analytischen Diffeomorphismus von $\mathbb{U}_L^{\mathbb{K}}$ auf $\mathbb{D}_L^{\mathbb{K}}$, dabei ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$, im Falle $\mathbb{K} = \mathbb{H}$ entsteht ein Diffeomorphismus zwischen $\mathbb{U}_L^{\mathbb{K}}$ und $I\mathbb{D}_L^{\mathbb{H}}$.

Beweis: Die Gültigkeit für $\mathbb{K} = \mathbb{C}$ wurde bereits in Satz 5.1 gezeigt. Im Falle $\mathbb{K} = \mathbb{H}$ ist zusätzlich zu beachten, dass

$$(IU)^*IU < \mathbb{1} \Leftrightarrow U^*I^*IU < \mathbb{1} \Leftrightarrow U^*U < \mathbb{1}$$

gilt. Damit kann der Beweis aus Satz 5.1 in weiten Teilen übernommen werden, und es ist nur noch zu überprüfen, dass die Symmetrie und die Schiefsymmetrie erhalten bleiben.

1. Zunächst zur Symmetrie, d. h. $\mathbb{K} = \mathbb{R}$. Sei $U \in \mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C})$, dann gilt:

$$\begin{aligned} \big(Z(U)\big)^T &= \big(i(\mathbb{1} - U)(\mathbb{1} + U)^{-1}\big)^T \\ &= i\big((\mathbb{1} + U)^{-1})\big)^T(\mathbb{1} - U)^T \\ &= i(\mathbb{1} + U)^{-1}(\mathbb{1} - U). \end{aligned}$$

Da schon im Beweis von Satz 5.1 gezeigt wurde, dass $\mathbb{1} + U$ und $\mathbb{1} - U$ kommutieren (und somit auch $(\mathbb{1} + U)^{-1}$ und $(\mathbb{1} - U)^{-1}$, folgt die Behauptung.

2. Ist andererseits $Z \in \mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C})$, so gilt:

$$\begin{aligned}\big(U(Z)\big)^T &= \big((Z-i\mathbb{1})(Z+i\mathbb{1})^{-1}\big)^T\\ &= (Z+i\mathbb{1})^{-1}(Z-i\mathbb{1})\\ &= (Z-i\mathbb{1})(Z+i\mathbb{1})^{-1}.\end{aligned}$$

   Die letzte Gleichheit folgt wie im ersten Fall aus dem Beweis von Satz 5.1.

3. Nun betrachtet man dan Fall $\mathbb{K} = \mathbb{H}$: Sei $Z \in \mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Skew}\,(L,\mathbb{C})$, dann folgt:

$$\begin{aligned}\big(U(Z)\big)^* &= \big((Z-i\mathbb{1})(Z+i\mathbb{1})^{-1}\big)^*\\ &= (-Z-i\mathbb{1})^{-1}(-Z+i\mathbb{1})\\ &= (Z-i\mathbb{1})(Z+i\mathbb{1})^{-1}.\end{aligned}$$

4. Weiter sei $U \in \mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C})$, dann gilt:

$$\begin{aligned}\big(Z(U)\big)^* &= \big(i(\mathbb{1}-U)(\mathbb{1}+U)^{-1}\big)^*\\ &= -i(\mathbb{1}+U)^{-1}(\mathbb{1}-U)\\ &= -\big(i(\mathbb{1}-U)(\mathbb{1}+U)^{-1}\big) = -\big(Z(U)\big).\end{aligned}$$

Daraus sind alle behaupteten Symmetrien gezeigt, und der Beweis ist abgeschlossen. □

Mit Hilfe von Satz B.3 kann man auch die Eigenschaften der Ränder des Einheitskreises bzw. der oberen Halbebene verallgemeinern.

Folgerung B.2 *Die Ränder von $\mathbb{U}_L^{\mathbb{R}}$ und $\mathbb{U}_L^{\mathbb{H}}$ sind wie im Falle $\mathbb{K} = \mathbb{C}$ stratifizierte Räume. Sie sind gegeben durch:*[4]

$$\begin{aligned}\partial\,\mathbb{U}_L^{\mathbb{R}} &= \partial\,\mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C}) &&\text{mit Strata } \partial_l\,\mathbb{U}_L^{\mathbb{R}} = \partial_l\,\mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C})\\ \partial\,\mathbb{U}_L^{\mathbb{H}} &= \partial\,\mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Skew}\,(L,\mathbb{C}) &&\text{mit Strata } \partial_l\,\mathbb{U}_L^{\mathbb{H}} = \partial_l\,\mathbb{U}_L^{\mathbb{C}} \cap \mathrm{Skew}\,(L,\mathbb{C}).\end{aligned}$$

Auf dieselbe Weise definiert man auch $\mathbb{U}_L^{\mathbb{R}}$ und $\mathbb{U}_L^{\mathbb{H}}$:

$$\begin{aligned}\partial\,\mathbb{D}_L^{\mathbb{R}} &= \partial\,\mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C}) &&\text{mit Strata } \partial_l\,\mathbb{D}_L^{\mathbb{R}} = \partial_l\,\mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Sym}\,(L,\mathbb{C})\\ \partial\,\mathbb{D}_L^{\mathbb{H}} &= \partial\,\mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C}) &&\text{mit Strata } \partial_l\,\mathbb{D}_L^{\mathbb{H}} = \partial_l\,\mathbb{D}_L^{\mathbb{C}} \cap \mathrm{Self}\,(L,\mathbb{C}).\end{aligned}$$

[4] Man beachte, dass man, falls $\mathbb{K} = \mathbb{H}$ ist, stets fordern muss, dass L gerade ist.

Der maximale Rand, der von besonderer Bedeutung ist, $\partial_L\, \mathbb{D}_L^{\mathbb{K}}$*, hat also die folgende Gestalt:*

$$\partial_L\, \mathbb{D}_L^{\mathbb{R}} = \mathfrak{U}(L) \cap \mathrm{Sym}\ (L, \mathbb{C}), \quad \partial_L\, \mathbb{D}_L^{\mathbb{H}} = \mathfrak{U}(L) \cap \mathrm{Self}\ (L, \mathbb{C}).$$

Daraus folgt nach Satz B.1, dass man auch für $\mathbb{K} = \mathbb{R}$ *oder* $\mathbb{H}$ *die maximalen Ränder des Einheitskreises mit den entsprechenden Lagrange-Ebenen identifizieren kann.*

In diesem Anhang wurde eine Verallgemeinerung diskutiert: Es stellte sich heraus, dass man auf der Basis der Aussagen über die Strukturen über den Körper $\mathbb{C}$ durch die Forderung nach zusätzlichen Symmetrien auch auf das Verhalten der jeweiligen Gruppen und Mannigfaltigkeiten über $\mathbb{R}$ und $\mathbb{H}$ schließen kann.

Anhang C

Auszüge aus der Theorie der linearen Differentialgleichungen

C.1 Existenz- und Eindeutigkeitssätze aus der Theorie der linearen Differentialgleichungen

Zunächst benötigt man die folgende

Definition C.1 *Sei* $U \subset \mathbb{R} \times \mathbb{C}^{2L \times L}$ *und* $f : U \to \mathbb{C}^{2L \times L}$, $(x, u_x) \mapsto f(x, u_x)$ *eine stetige Abbildung. Dann genügt* f *einer* lokalen Lipschitz-Bedingung, *falls für jedes* $(x_0, u_{x_0}) \in U$ *eine offene Umgebung* $U_\delta \subset U$ *und eine Zahl* $C > 0$ *existieren, so dass*

$$\|f(x, u_x) - f(x, \widetilde{u_x})\| \leq C \; \|u_x, \widetilde{u_x}\|$$

für alle (x, u_x), $(x, \widetilde{u_x}) \in U_\delta$ *gilt.*

Der nächsten Bemerkung stelle ich noch einen Hinweis zur Notationen voran: Sei $u_x \in \mathbb{C}^{2L}$. Dann bezeichne ich die n-te Zeile von u_x ($n = 1, \ldots, 2L$) mit u_x^n.

Bemerkung C.1 *Ist* $f(x_0, u_x^1, \ldots, u_x^{2L})$ *in* $u_x^1, \ldots, u_x^{2L}$ *stetig differenzierbar, dann genügt* f *einer lokalen Lipschitz-Bedingung.*

Denn (vgl. [Mu03, Kap. 23, S. 227]): Es sei (x_0, u_{x_0}) in U. Dann gibt es ein U_δ mit $\overline{U_\delta} \subset U$. Nach Voraussetzung ist $\frac{\partial f}{\partial u_x^n} : U \to \mathbb{C}^{2L \times 2L}$ stetig (für alle $n = 1, \ldots, 2L$).

Da $\overline{U_\delta}$ kompakt ist, existiert

$$C^n := C^n(U) := \max_{(x,u_x^n)\in\overline{U_\delta}} \left\| \frac{\partial f}{\partial u_x^n} \right\|.$$

Wegen der Endlichkeit von n folgt damit sofort die Existenz des Maximums für alle u_x^n:

$$C := C(U) := \max_{n=1,\ldots,2L} |C^n(U)|.$$

Da die Menge $\left\{u_x \in \mathbb{C}^{2L\times L} : (x, u_x) \in U\right\}$ konvex ist, gilt:

$$|f(x, u_x) - f(x, \widetilde{u_x})| \leq C\,|u_x - \widetilde{u_x}|. \qquad \square$$

Ganz allgemein gilt:

Satz C.1 (Existenz- und Eindeutigkeitssatz von Picard-Lindelöf) Lokale Version
Sei $u_y' = f(x, u_x)$ eine Differentialgleichung, die einer lokalen Lipschitz-Bedingung genügt[1]*. Dann gibt es zu jedem $(x_0, u_{x_0}) \in U$ ein $\delta > 0$ und genau eine Lösung $\varphi : (x_0 - \delta, x_0 + \delta) \to \mathbb{C}^{2L\times L}$ der Differentialgleichung $u_x' = f(x, u_x)$, die die Randbedingung $\varphi(x_0) = u_{x_0}$ erfüllt.*

Beweise findet man etwa bei Arnold' [Ar01, Kap. 4, S. 275–281] oder in [Mu03, Kap. 23, S. 229–231].

Mit Hilfe dieses Satzes wird die Existenz einer Lösung nur auf einer (evtl. sehr kleinen) Umgebung sicher gestellt. Daher stelle ich noch eine globale Version des Satzes vor. Dazu benötigt man aber zunächst noch die folgende Definition.

Definition C.2 *Seien (φ, I) und $(\widetilde{\varphi}, \widetilde{I})$ zwei Lösungen des Anfangswertproblems $u_x' = f(x, u_x)$ auf Intervallen I bzw. $\widetilde{I}$. Dann heißt $(\widetilde{\varphi}, \widetilde{I})$ eine* Fortsetzung *von (φ, I), falls $I \subset \widetilde{I}$ ist und $\widetilde{\varphi}\big|_I = \varphi$ gilt. Eine Lösung (φ_{max}, I_{max}) heißt* maximal, *wenn sie Fortsetzung jeder Lösung (φ, I) ist.*

Satz C.2 (Existenz- und Eindeutigkeitssatz von Picard-Lindelöf) Globale Version
Es sei $D \subset \mathbb{R}\times\mathbb{C}^{2L}$. Ferner sei $u_x' = f(x, u_x)$ eine Differentialgleichung, wobei $f \in$

[1] Man kann auch zeigen, dass auch ohne die Voraussetzung einer lokalen Lipschitz-Bedingung die Existenz einer Lösung auf einer Umgebung von (x_0, u_{x_0}) gesichert ist (Existenzsatz von Cauchy-Peano, siehe [CL55, Kap. 1.1, S. 6/7]. Allerdings sind diese Lösungen im Allgemeinen nicht eindeutig.

$C(D, \mathbb{C}^{2L})$ auf D einer lokalen Lipschitz-Bedingung genüge. Dann gibt es zu jedem $(x_0, u_{x_0}) \in D$ genau eine maximale Lösung (φ_{max}, I_{max}) des Anfangswertproblems. Weiterhin „verlässt" (φ_{max}, I_{max}) jede kompakte Teilmenge von D, d. h. zu jedem kompakten D existieren $X_1, X_2 \in I_{max}$ mit $(x, \varphi_{max}(x)) \notin K$ für alle $x < X_1$ und $x > X_2$.

Ein **Beweis** dazu findet sich in [Mu03, Kap. 23, S. 232–234].

Satz C.3 (Abhängigkeit der Lösungen von einem Parameter)
Um darüber hinaus die Abhängigkeit *eines Systems $u'_x = f(x, u_x, E)$, $u_x \in \mathbb{C}^{2L}$, $x \in I$ mit Anfangsbedingung $u_0 = \xi$* von einem Parameter E *zu untersuchen, führt man diesen Parameter E als fiktive Variable in die ursprünglichen Gleichungen ein. Es sei*

$$\widetilde{u}_x \in \mathbb{C}^{2L+1}, \quad (\widetilde{u}_x)_i = (u_x)_i \ (i = 1, \ ..., \ 2L), \quad (\widetilde{u}_x)_{2L+1} = E.$$

Man betrachtet also das System

$$\widetilde{u}'_x = \widetilde{f}(x, \widetilde{u}_x) \qquad (x \in I)$$

mit

$$\widetilde{f}_i(x, \widetilde{u}_x) = f_i(x, u_x, E) \ (i = 1, \ ..., \ 2L), \qquad \widetilde{f}_{2L+1}(x, \widetilde{u}_x) = 0,$$

und unterstellt, dass f_i eine analytische Funktion in E ist $(i = 1, \ \ldots, \ 2L)$. Die Anfangsbedingung wird damit zu:

$$(\widetilde{u}_0)_i = \xi_i \ (i = 1, \ ..., \ 2L), \qquad (\widetilde{u}_0)_{2L+1} = E.$$

Die Lösung dieses Anfangswertproblems ist gegeben durch eine Funktion $\varphi = \varphi(x, u_x, \xi, E)$. Aus dem Satz von Picard-Lindelöf folgt die Existenz einer Lösung auf gewissen Umgebungen U, falls f einer lokalen Lipschitz-Bedingung genügt. Außerdem ist φ analytisch in allen Variablen, also auch in E.

Einen detaillierten **Beweis** dazu liefert [CL55, Kap. 1, S. 30/31].

C.2 Methoden zur Lösung von Differentialgleichungen mit konstanten Koeffizienten

Wie man die Lösungen von Differentialgleichungen mit konstanten Koeffizienten bestimmt, ist weitgehend bekannt. Die für diese Arbeit gewählte Darstellung sttzt sich auf ([Mu03], Kap. 25), aber die hier genutzten Sätze finden sich beispielsweise auch in [Am95], [Ar01], [CL55] oder [Wa93].

Vorweg zur Definition der Matrixexponentialfunktionen:

Definition C.3 *Es sei* $A \in \mathrm{Mat}(2L \times 2L,\, \mathbb{C})$ *Dann definiert man die Matrixexponentialfunktion als*

$$e^{A\,x} = \sum_{k=1}^{\infty} \frac{(A\,x)^k}{k!}.$$

Bemerkung C.2 *Aus Definition C.3 ergibt sich sofort:*[2]

$$\left(e^{A\,x}\right)' = A\,e^{A\,x}. \tag{C.1}$$

Außerdem gilt[3] *für jede invertierbare Matrix* $\mathfrak{M} \in \mathrm{Mat}(2L \times 2L,\, \mathbb{C})$*:*

$$e^{\mathfrak{M}^{-1}\,A\,\mathfrak{M}} = \mathfrak{M}^{-1}\,e^{A}\,\mathfrak{M}. \tag{C.2}$$

Aus diesen Bemerkungen ergibt sich folgender Satz:

Satz C.4 *Es sei* $A \in \mathrm{Mat}(2L \times 2L,\, \mathbb{C})$*. Dann ist für alle* $u_0 \in \mathbb{C}^{2L}$ *eine Lösung des Anfangswertproblems*

$$u_y' = A\,u_y \qquad \text{und} \qquad u_y(0) = u_0$$

gegeben durch

$$u_y(x) = e^{A\,x}\,u_0.$$

Beweis (vgl. [CL55, Kap. 3, S. 75/76]): Mit Hilfe der Produktregel (C.1) berechnet man:

$$\left(e^{A\,x}\,u_0\right)' = \left(e^{A\,x}\right)'\,u_0 = Ae^{A\,x}\,u_0,$$

[2] Diese Behauptung folgt, wenn man die Exponentialfunktion als Reihe ausschreibt und dann ausnutzt, dass der Konvergenzradius dieser Reihe ∞ ist. Das bedeutet, man kann die Limiten vertauschen und somit die Ableitung unter die Summe schreiben. Daraus ergibt sich sofort die Behauptung (siehe [Mu03, Kap. 25, S. 254–256]).

[3] Der Beweis dazu folgt per Induktion aus der Reihenentwicklung, wenn man beachtet, dass gilt:

$$(\mathfrak{M}^{-1}\,A\,\mathfrak{M})^k = \mathfrak{M}^{-1}\,A^k\,\mathfrak{M},$$

für den explizit ausgefhrten Beweis sei wiederum verwiesen auf [Mu03, Kap. 25, S. 257/258].

d. h. $x \mapsto e^{A\,x}\, u_0$ ist eine Lösung der Differentialgleichung. Außerdem wird die Anfangsbedingung erfüllt, denn es gilt:

$$e^{A\,x}\, u_0 \Big|_{x=0} = \mathbb{1}_{2L}\, u_0 = u_0. \qquad \square$$

Zur konkreten Berechnung von Lösungen bedient man sich häufig der folgenden Formel, die auf (C.2) basiert: Sei $B \in \mathrm{Mat}(L \times L,\, \mathbb{C})$ eine diagonalisierbare Matrix und sei $\mathfrak{M}$ ein Basiswechsel, der $B = \mathfrak{M}\,\Lambda\,\mathfrak{M}^{-1}$ erfüllt. Dabei sei $\Lambda = \mathrm{diag}\,(\lambda_1,\ \ldots,\ \lambda_L)$. Dann gilt:

$$e^{B\,x} = e^{\mathfrak{M}\,\Lambda\,\mathfrak{M}^{-1}} \overset{(C.2)}{=} \mathfrak{M}\, e^{\Lambda\,x}\, \mathfrak{M}^{-1} = \mathfrak{M}\, \mathrm{diag}\, \left(e^{\lambda_1\,x},\ \ldots,\ e^{\lambda_L\,x}\right) \mathfrak{M}^{-1}.$$

Danke!

Ein mathematisches Problem sei ferner schwierig, damit es uns reizt;
und dennoch nicht völlig unzugänglich, damit es unserer Anstrengung nicht spotte;
es sei uns ein Wahrzeichen auf den verschlungenen Pfaden zu verborgenen Wahrheiten
– uns hernach lohnend mit Freude über die gelungene Lösung.

David Hilbert, „Mathematische Probleme“
Vortrag beim Internationalen Mathematischen Kongress in Paris, 1900

An dieser Stelle möchte ich allen danken, die dazu beigetragen haben, dass ich nun tatsächlich „mit Freude“ auf diese Diplomarbeit schauen kann.

An erster Stelle steht mein Betreuer Prof. Dr. Hermann Schulz-Baldes, dem ich für seine Unterstützung und seinen fachlichen Rat danken möchte. Christian Sadel danke ich für die anregenden Diskussionen.

Herzlich bedanken möchte ich mich auch bei meinem Freund Philipp Welker, der nicht nur korrigierend den Rotstift schwang, sondern mir auch stets ein großer Rückhalt war. Simone Heinrich danke ich für die vielen lustigen Stunden in der WG, für die Versorgung mit Schokolade und für ihre Ratschläge. Auch meinen Freunden aus der Zeit in Trier, Lille und Erlangen und den Kollegen aus dem Michael-Müller-Verlag gilt mein Dank. Ihr habt meine Studienzeit geprägt und bereichert. Da mir das Durchhalten ohne meine tägliche Koffeindosis schwer gefallen wäre, danke ich auch allen Kaffeebauern dieser Welt.

Last but not least möchte ich mich bei meinen Eltern bedanken, die mir mein Studium ermöglicht und mich in all meinen Entscheidungen unterstützt haben.

Literaturverzeichnis

[Am95] Amann, Herbert: *Gewöhnliche Differentialgleichungen.* W. de Gruyter, Berlin, 2. Auflage 1995.

[Ar67] Arnol'd, Vladimir I.: *Charakteristic Class entering in quantization conditions.* Funct. Analysis Applic. **1**, S. 1–13, 1967.

[Ar88] Arnol'd, Vladimir I.: *Mathematical Methods of Classical Mechanics.* Springer, New York, 2. Auflage 1988.

[Ar85] Arnol'd, Vladimir I.: *The Sturm Theorems and Symplectic Geometry.* Func. Anal. Appl. **19,**(4), S. 1–10, 1985.

[Ar00] Arnol'd, Vladimir I.: *The Complex Lagrangian Grassmannian.* Funct. Anal. Applic. **34,** S. 208–210, 2000.

[Ar01] Arnol'd, Vladimir I.: *Gewöhnliche Differentialgleichungen.* Springer, Berlin, 2. Auflage 2001.

[BE05] Bennewitz, Christer und Everitt, W. Norrie: *The Titchmarsh-Weyl Eigenfunction Expansion Theorem for Sturm-Liouville Differential Equations.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 137–171, Birkhäuser Verlag, Basel, 2005.

[Be98] Berndt, Rolf: *Einführung in die Symplektische Geometrie.* Vieweg, Wiesbaden, 1998.

[Bo04] Bosch, Siegfried: *Algebra.* Springer, Berlin, 5. Auflage 2004.

[Bo56] Bott, Raoul: *On the iteration of Closed Geodesics and the Sturm Intersection Theory.* Commun. Pure Appl. Math. **9,** S. 171–206, 1956.

[CG01] Clark, Steve und Gesztesy, Fritz: *Weyl-Titchmarsh M-Function asymptotics for matrix valued Schrödinger Operators.* Proc. London Math. Soc. **82,** S. 701–724, 2001.

[CGIP03] Contreras, Gonzales, Gambaudo, Jean-Marc, Iturriaga, Renato und Paternain, Gabriel: *The asymptotic Maslov Index and its applications.* Ergodic Theory and Dynamical Systems 2003, **23:** S. 1415–1443.

[CL55] Coddington, Earl A. und Levinson, Norman: *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, New York, 1955.

[CLR90] Creagh, Stephen C., Littleton, Robert G. und Robbins, Jonathan M.: *Geometrical properties of Maslov indices in the semiclassical trace formula for the density of states.* Physical Review A, **42,** 4, S. 1907–1922, 1990.

[deG97] de Gosson, Maurice: *Maslov Classes, Metaplectic Representation and Lagrangian Quantization.* Akademie Verlag, Berlin, 1997.

[deG06] de Gosson, Maurice: *Symplectic Geometry and Quantum Mechanics.* Birkhäuser Verlag, Basel, 2006.

[deR05] del Río, Rafael: *Boundary Conditions and Spectra of Sturm-Liouville Operators.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 217–235, Birkhäuser Verlag, Basel, 2005.

[Ev05a] Everitt, W. Norrie: *Charles Sturm and the Development of Sturm-Liouville Theory in the Years 1900 to 1950.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 45–74, Birkhäuser Verlag, Basel, 2005.

[Ev05b] Everitt, W. Norrie: *A Catalogue of Sturm-Liouville Differential Expressions.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 271–332, Birkhäuser Verlag, Basel, 2005.

[Fi00] Fischer, Gerd: *Lineare Algebra.* Vieweg, Braunschweig/Wiesbaden, 12. Auflage, 2000.

[FJN03] Fabbri, Roberta, Johnson, Russell und Núñez, Carmen: *Rotation number for non-autonomous linear Hamiltonian systems I: Basic properties.* Zeitschrift für angewandte Mathematik und Physik, **54,** S. 484–502, 2003.

[Ga94] Gardner, Martin: *Gardner on Gardner.* JPBM Communications Award Presentation. Focus – The Newsletter of the Mathematical Association of America, **14,** 6, 1994.

[Go77] Goering, Hermann: *Asymptotische Methoden zur Lösung von Differentialgleichungen.* Vieweg, Braunschweig, 1. Auflage, 1977.

[Ha57] Halmos, P. R.: *Nicholas Bourbaki.* Scientific American, S. 88/89, Mai 1957.

[Hil64] Hilbert, David: *Über das Unendliche* (1925). In: Hilbertiana. Fünf Aufsätze von David Hilbert. S. 1–28, Darmstadt, 1964.

[Hin05] Hinton, Don: *Sturm's 1836 Oscillation Results. Evolution of the Theory.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 45–74, Birkhäuser Verlag, Basel, 2005.

[HS96] Hislop, Peter D. und Sigal, Israel Michael: *Introduction to Spectral Theory. With Applications to Schrödinger Operators.* Springer, New York, 1996.

[JM82] Johnson, R. und Moser, J.: *The Rotation Number for Almost Periodic Potentials.* Commun. Math. Phys. **42,** S. 403–438, 1982.

[Mu03] Müller, Jürgen: *Analysis I–IV. Skriptum zur Vorlesung Wintersemester 2001/2002 bis Sommersemester 2003.* Vorlesungsskript, www.mathematik.uni-trier.de/˜mueller/AnalysisI-IV.pdf.

[Na03] Nakahara, Mikio: *Geometry, Topology and Physics.* Taylor & Francis, New York, 2. Auflage 2003.

[NNO98] Novo, Sylvia, Nuñez, Carmen und Obaya, Rafael: *Ergodic properties and rotation number for linear Hamiltonian Systems.* J. Differential Equations, **148,** S. 148–185, 1998.

[Pe88] Pearson, D. B.: *Quantum Scattering and Spectral Theory.* Academic Press, London, 1988.

[Ro92] Robbins, Jonathan: *Winding number formula for Maslov indices.* CHAOS **2,** (1), 1992.

[Sa99] Sakhnovich, Lev A.: *Spectral Theory of Canonical Differential Systems. Method of Operator Identities.* Birkhäuser, Basel, 1999.

[SB06] Schulz-Baldes, Hermann: *Rotation numbers for Jacobi matrices with matrix entries.* 2006, Preprint, arXiv:math-ph/0702050.

[Si71/73] Siegel, Carl Ludwig: *Topics in complex function theory. Band II und III.* Wiley Interscience, New York 1971 und 1973.

[Wa93] Walter, Wolfgang: *Gewöhnliche Differentialgleichungen. Eine Einführung.* Springer, Berlin, 5. Auflage 1993.

[Wei76] Weidmann, Joachim: *Lineare Operatoren in Hilberträumen.* B. G. Teubner, Stuttgart, 1976.

[Wei87] Weidmann, Joachim: *Spectral Theory of Ordinary Differential Operators. Lecture Notes in Mathematics.* Springer, Berlin, 1987.

[Wei05] Weidmann, Joachim: *Spectral Theory of Sturm-Liouville Operators. Approximation by Regular Problems.* In: Amrein, Werner O., Hinz, Andreas M., Pearson, David B. (Hgs.): Sturm-Liouville Theory. Past and Present. S. 75–98, Birkhäuser Verlag, Basel, 2005.

[Wer04] Werner, Dirk: *Funktionalanalysis.* Springer, Berlin, 5. Auflage 2004.

[Wey10] Weyl, Hermann: *Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen.* Math. Annalen **68,** S. 220–269, 1910.

[Ze05] Zettl, Anton: *Sturm-Liouville Theory.* Mathematical surveys and monographs; no. 121. American Mathematical Society, Providence, 2005.

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Erlangen, 26. Juli 2007

Martina Comes